Deliverables, Previous and Future Work

Joaquín Rapela and Maneesh Sahani

Gatsby Computational Neuroscience Unit

February 18, 2025

Contents

Contents

Aims

1. share expertise between the SWC/GCNU and the AIND in NaLoDuCo experimentation to jointly advance the field.
2. disseminate hardware and software technology for NaLoDuCo experimentation to the broad research community.

Note

Hardware and software generated in this collaboration will be evaluated in three different experimental setups:

1. freely-moving foraging mice (SWC)
2. head-fixed foraging mice (AIND)
3. freely-moving odour exploration mice (AIND)

Contents

Deliverables

1. hardware specifications for recordings of behaviour and neural activity used at the SWC/GCNU[1] and at the AIND for head-fixed foraging[2] and freely moving odour exploration.
2. software for managing long-duration recordings (e.g., data storage, data indexing).
3. software for online/offline quality control.
4. software for creating alerts.
5. software for online (behavioural and neural) data visualisation.
6. software for online (behavioural and neural) data analysis.

---

[1] https://sainsburywellcomecentre.github.io/aeon_docs/reference/hardware.html

[2] https://www.allenneuraldynamics.org/platforms/behavior

Previous work

- the SWC has performed foraging experiments
  - lasting 2 weeks and recording behaviour only
  - lasting 27 hours and recording behaviour and electrophysiology
  - data is stored in files and in a MySQL database
- the AIND has performed foraging experiments in head-fixed mice. These experiments are a few hours long.
- the AIND is setting up odour exploration experiments that will last several days.
- items 1–2: above have been completed for the SWC foraging experiments
- item 3: the SWC has developed some quality control software for behaviour.
- item 4: alerts have been developed for behaviour and electrophysiology.

Previous work

- item 5: the SWC has developed some online behavioural data visualisation tools in Bonsai.
- item 6: funded by BBSRC, we have integrated into Bonsai tools for online data analysis:
  - estimate kinematics of mice
  - estimate kinematic states of mice using Hidden Markov Models
  - clusterless point-process decoder of mice position and replay from electrophysiology recordings.
- disseminated documentation of hardware used at the AIND to perform head-fixed foraging experiments[3] in virtual reality setups.

[3] https://www.allenneuraldynamics.org/platforms/behavior

Previous work

- disseminated documentation on software used at the SWC to control NaLoDuCo foraging experiments (see repo)
- disseminated documentation on machine learning methods integrated into Bonsai for analysing behavioural and neural time series in real time (see repo)

Future work

- item 3: the SWC and the AIND have developed several tools for offline quality control. Next, we need to build online versions of them.
- item 4: develop more software for data visualisation.
- items 5 and 6: develop more software for online data analysis.
  - online estimate of latent variables from Neuropixels recordings.
  - online estimate of RL models.
- assist Dr. Carl Schoonover (AIND) on the use hardware and software developed at the SWC/GCNU to create olfactory learning NaLoDuCo experiments.

Contents

Deliverables

1. dashboard to convert data collected in NaLoDuCo experiments to the Zarr format, and to upload the Zarr files to DANDI.
2. dashboard to stream in real time to DANDI data collected from NaLoDuCo experiments

Previous work

- the AIND is experienced on sharing their recordings on DANDI. However, these recordings are not as large as those in NaLoDuCo experiments and they are not continual.

Future work

- develop dashboard to convert data collected in NaLoDuCo experiments to the Zarr format
- develop dashboard to stream in real time data collected from NaLoDuCo experiments to DANDI
- DANDI is typically used to store neurophysiological datasets much smaller than those generated in NaLoDuCo experiments. Conventional methods to access data in DANDI may not be fast enough to allow performant data visualisation and/or data analysis. We may need to explore parallel computing and/or resource efficient cloud configurations (i.e., optimise cloud configurations to improve runtime performance).

Contents

Deliverables

1. web-based dashboard for **online** experiment monitoring.
2. web-based dashboard for **online** data analysis and visualisation of its results.
3. web-based dashboard for **offline** visualisation of NaLoDuCo behavioural and neural recordings on DANDI. These dashboards should allow users to efficiently **visually explore** very large behavioural, neural and data analysis datasets. For instance, users should be able to quickly visualise videos of mice when they were next to any location of the arena, at any time of the day, and in any behavioural state (as inferred by a Hidden Markov Model). These efficient data explorations will require sophisticated data indexing schemes.
4. web-based dashboard for **offline** visualisation of data analysis results on DANDI.

Previous work

- Neurosift (repo, paper) allows to visualise shorter-duration behavioural and neural recordings in DANDI.
- Dendro (repo) allows to perform analysis on the cloud and visualise the results of such analysis
- the SWC has developed visualisations in Bonsai for experimental monitoring.
- offline and precomputed visualisations developed at the SWC, of some experimental variables (e.g., total distance that animals have moved the wheel, proportion of time animals spent on the nest, corridor or food patches).
- offline and precomputed visualisations developed at the AIND for shorter-duration experiments.
- offline and precomputed visualisations from IBL for short-duration experiments.

Future work

- create web-based dashboard to monitor NaLoDuCo experiments.
- create web-based dashboard for **online** data analysis and visualisation of its results.
- create web-based dashboard for **offline** visualisation of NaLoDuCo behavioural and neural recordings on DANDI. Neurosift has been designed to visualise relatively short duration datasets. We will extend it with data pyramids (e.g., https://github.com/carbonplan/ndpyramid) to enable it to operate on long-duration recordings.
- web-based dashboard for **offline** visualisation of data analysis results on DANDI.
  Dendro has been used to analyse on the cloud relatively small datasets. We will extend Dendro to analyse very large datasets and visualise the results of these analysis.

Contents

Deliverables

1. methods (and software implementations) for sorting spikes from Neuropixels probes in long-duration experiments
2. software for curation of results of sorting spikes from Neuropixels probes in long-duration experiments
3. quality control measures for the results of sorting spikes from Neuropixels probes in long-duration experiments

Previous work

- methods exist for sorting, curating and quality control spikes from short duration experiments
- Dr. Carl Schoonover (AIND) has developed methods to sort spikes from long-duration experiments (find out more about these methods)
- the SWC has managed to sort spikes of long-duration experiments, from a small subset of channels of a Neuropixels probe (how many channels? what duration?)

Contents

Deliverables

1. methods to analyse behavioural and electrophysiological recordings from NaLoDuCo experiments.
2. integration of these methods into DANDI, as in Dendro, so that users can run them on NaLoDuCo datasets stored in DANDI.

Previous work

1. At the Gatsby we have pioneered several methods for the characterisation of neural time series (e.g., Yu et al., 2009; Duncker and Sahani, 2018; Rutten et al., 2020; Yu et al., 2024; Buesing et al., 2012b,a; Macke et al., 2015; Soulat et al., 2021; Walker et al., 2023; Turner and Sahani, 2014; O'Shea et al., 2022; Pachitariu et al., 2013a,b)
2. Dendro allows to perform advanced data analysis on DANDI, and to reuse previous analysis.

Future work

1. distribute methods for the analysis of behavioral and electrophysiological NaLoDuCo recordings. Consider needs to develop online and/or adaptive versions of these methods.

   For behavioural data, we will investigate methods to

   - track multiple body parts of animals (deep neural networks)
   - infer kinematics of foraging mice (linear dynamical systems)
   - segment behaviour into discrete states
   - characterise short- and long-term periodicities in behaviour
   - infer the rules that govern mice behaviour from behavioural observations only] (i.e., policy inference).

   For neural data, we will investigate methods to:

   - estimate low-dimensional continual representations of neural activity (i.e., latents inference)
   - segment neural activity into discrete states
   - characterise short- and long-term periodicities in neural activity
   - decode environment variables from neural activity

Contents

Deliverables

1. software, hardware and papers demonstrating applications of intelligent experimental control in neuroscience experimentation.

Previous work

- Bonsai is an excellent software for close-loop neuroscience experimental control.
- funded by a BBSRC award[3] we are integrating machine learning methods into Bonsai.

[3] https://gow.bbsrc.ukri.org/grants/AwardDetails.aspx?FundingReference=BB%2FW019132%2F1

Future work

- Bonsai has been used for close-loop control using direct observations. With the advanced machine learning methods that we have added to Bonsai, we can know infer subtle patterns in behavioural and neural recordings not visible to the naked eye. Now we will find applications where these patterns can be used for experimental control. Fortunately, we work at the SWC, where such applications abound.
- it would be very useful to perform control based on patterns inferred from spiking activity recorded by Neuropixels probes. To do so, we need online spike sorting methods for Neuropixels probes. These methods are underdeveloped and, if necessary, we will need to develop one.

References I

Buesing, L., Macke, J. H., and Sahani, M. (2012a). Learning stable, regularised latent models of neural population dynamics. *Network: Computation in Neural Systems*, 23(1-2):24–47.

Buesing, L., Macke, J. H., and Sahani, M. (2012b). Spectral learning of linear dynamics from generalised-linear observations with application to neural population data. In *Advances in Neural Information Processing Systems 25: 26th Conference on Neural Information Processing Systems (NIPS 2012)*, pages 1691–1699.

Duncker, L. and Sahani, M. (2018). Temporal alignment and latent gaussian process factor inference in population spike trains. In *Advances in Neural Information Processing Systems*, pages 10445–10455.

References II

Macke, J. H., Buesing, L., and Sahani, M. (2015). *Estimating state and parameters in state space models of spike trains*, page 137–159. Cambridge University Press.

O'Shea, D. J., Duncker, L., Goo, W., Sun, X., Vyas, S., Trautmann, E. M., Diester, I., Ramakrishnan, C., Deisseroth, K., Sahani, M., et al. (2022). Direct neural perturbations reveal a dynamical mechanism for robust computation. *bioRxiv*, pages 2022–12.

Pachitariu, M., Packer, A. M., Pettit, N., Dalgleish, H., Hausser, M., and Sahani, M. (2013a). Extracting regions of interest from biological images with convolutional sparse block coding. *Advances in neural information processing systems*, 26.

References III

Pachitariu, M., Petreska, B., and Sahani, M. (2013b). Recurrent linear models of simultaneously-recorded neural populations. *Advances in neural information processing systems*, 26.

Rutten, V., Bernacchia, A., Sahani, M., and Hennequin, G. (2020). Non-reversible gaussian processes for identifying latent dynamical structure in neural data. *Advances in neural information processing systems*, 33:9622–9632.

Soulat, H., Keshavarzi, S., Margrie, T. W., and Sahani, M. (2021). Probabilistic tensor decomposition of neural population spiking activity. In *Advances in Neural Information Processing Systems*.

Turner, R. E. and Sahani, M. (2014). Time-frequency analysis as probabilistic inference. *IEEE Transactions on Signal Processing*, 62(23):6171–6183.

References IV

Walker, W. I., Soulat, H., Yu, C., and Sahani, M. (2023). Unsupervised representation learning with recognition-parametrised probabilistic models. In *International Conference on Artificial Intelligence and Statistics*, pages 4209–4230. PMLR.

Yu, B. M., Cunningham, J. P., Santhanam, G., Ryu, S. I., Shenoy, K. V., and Sahani, M. (2009). Gaussian-process factor analysis for low-dimensional single-trial analysis of neural population activity. *Journal of neurophysiology*, 102(1):614–635.

Yu, C., Sahani, M., and Lengyel, M. (2024). Discovering temporally compositional neural manifolds with switching infinite gpfa. *bioRxiv*, pages 2024–10.