BRAIN–COMPUTER INTERFACES

# From unstable input to robust output

In the presence of recording instabilities, the performance of brain–computer interfaces can be robustly maintained by exploiting 'hidden' structures underlying neural activity.

Lahiru N. Wimalasena, Lee E. Miller and Chethan Pandarinath

In people with paralysis, brain–computer interfaces (BCIs) can restore voluntary movements by interfacing directly with the brain to translate movement intention into action. The best-performing BCIs monitor movement-related neural signals via implanted microelectrodes. To translate the monitored signals into commands, a decoder is trained to find a mapping from recorded neural activity to a control signal. Progress in the development of BCIs has enabled their use in a range of applications, such as rapid typing, the control of anthropomorphic robotic arms, the production of synthetic speech and the stimulation of paralysed muscles to enable reaching and grasping[1–4]. However, neural recording instabilities incurred over time present challenges to maintaining robust closed-loop performance. For example, slight displacements of the implanted electrodes (relative to the surrounding brain tissue) can cause changes in recorded neuron identity and lead to intraday and inter-day instabilities, confounding the decoding of intent[5,6]. Reporting in *Nature Biomedical Engineering*, Byron Yu and colleagues now show that the decoding performance of BCIs can be stabilized by harnessing 'hidden' structures (known as low-dimensional neural manifolds) underlying the activity of large numbers of neurons[7].

Neural manifolds represent patterns of coordinated activity across neurons that would be unidentifiable by looking only at individual neuronal activity[8,9] (Fig. 1a). They are thought to reflect constraints imposed by the underlying neural circuitry[9]. BCI decoders that rely on neural manifolds use a two-stage approach: a dimensionality-reduction stage to map the activity of individual neurons onto the underlying manifold, followed by the mapping of the manifold onto movements. And because manifolds are computed from a small, random sample of the cortical neurons, many different sets of recorded neurons can be mapped onto the same manifold[10–14]. These manifolds and their decoded output have a consistent relationship with behaviour

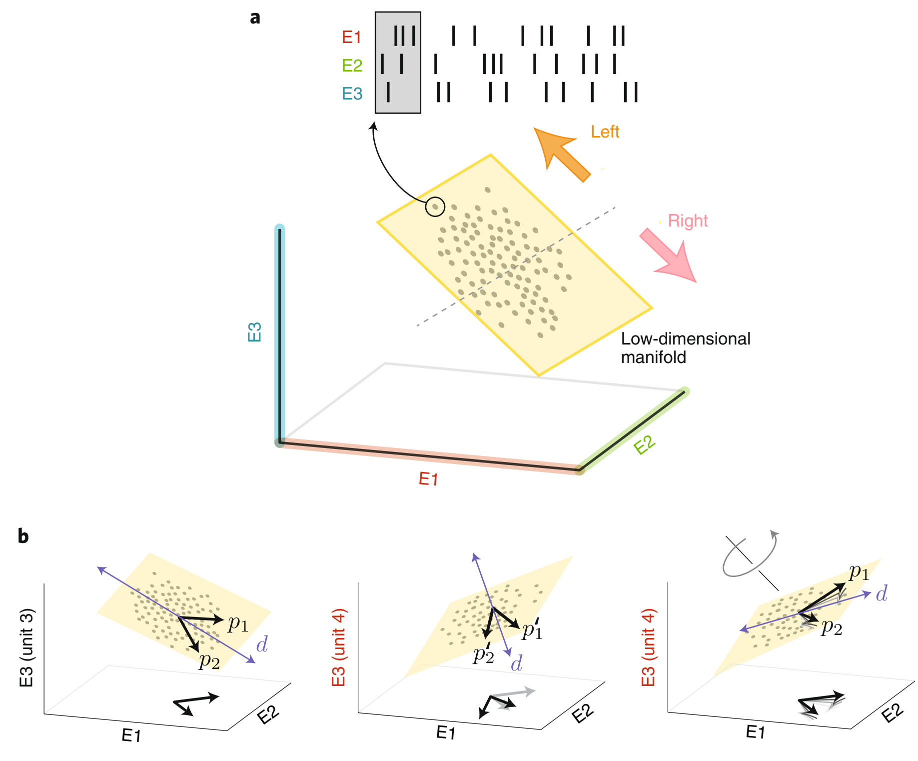


**Fig. 1 | Patterns in the underlying activity of simultaneously recorded neurons can be leveraged to improve the robustness of decoders for BCIs. a**, The spiking activity of neurons (vertical marks at the top of the schematic) at any point in time can be represented as a point in a 'state space' (with as many dimensions as electrodes), where each axis represents the degree of activity recorded by a single electrode. Not all possible patterns of activity are observed; rather, neural activity is confined to a lower-dimensional manifold (in this example, a two-dimensional plane within the three-dimensional state space). To decode the movements intended by the user of the BCI, a manifold-decoding approach projects neural activity onto a set of one or more output axes (the arrows here denote a single axis). **b**, Left: a low-dimensional manifold, which can be obtained via factor analysis, can be defined by a set of vectors (*p*) within the higher-dimensional state space of neural activity. The decoder's output relies on a set of decoding axes (*d*) defined in reference to *p* (for simplicity, only a single axis is shown). Middle: during the use of the BCI, new estimates of the manifold (*p*′) are periodically made. Right: accurate decoding from the new manifold requires finding an orthogonal transformation that orients *p*′ so as to map onto *p*. This involves the identification of 'reference' electrodes (black) that remained stable with respect to the manifold. *p*′ differs from *p* in its relationship to individual electrodes because of recording instabilities, as some electrodes no longer monitor the same neurons (red). Panel **b** adapted with permission from ref. [7], Springer Nature Limited.

across timescales ranging from months to years[10,12,14]. Hence, stable decoding can be achieved by properly recalibrating the dimensionality-reduction algorithm to correctly map the new set of neurons onto the same manifold[10–14]. This can be performed without supervision[11,13].

Two complementary strategies represent the leading methods to reduce the reliance on supervised recalibration (the need to collect a new labelled training dataset to correct for neural signal instability, which interrupts device use, increases training time and user effort, and needs the assistance of a technician or caregiver): training neural network decoders by using month-long datasets that expose the decoder to a wide variety of neurons and thus limits the amount of human supervision required[5], and the automatic adaptation of decoders without the need for explicit supervised recalibration[15]. The use of large training datasets and neural network decoders can achieve impressive robustness, as shown in a two-dimensional cursor movement task performed by non-human primates[5]. However, collecting such large supervised datasets would require a substantial time commitment from the user and is therefore challenging to perform clinically. Instead, semi-supervised approaches recalibrate decoders by using the retrospective analysis of data collected during the subject's normal use of the BCI[8]. For instance, in a setting with predefined targets, the neural activity preceding movement to a given target likely reflects the subject's intention to move toward that target. This knowledge can be used to update the decoder (much like the fully supervised case). For clinical BCIs, semi-supervised recalibration is state-of-the-art and confers stability across months of BCI use. However, it works only when the user's intent can be guessed post-hoc (as in BCI spellers or movement among a limited number of predefined targets). Thus, it is unlikely to scale to complex, naturalistic settings, such as the operation of a prosthetic robotic arm with several degrees of freedom moving without constraints.

Yu and co-authors' method stabilizes manifold decoders during online use by periodically updating the dimensionality-reduction stage in an unsupervised fashion (Fig. 1b). Their dimensionality reduction uses factor analysis, a statistical method for finding unobserved variables that explain relations among a larger number of correlated, observed variables. The authors' approach iteratively aligns the estimated manifold to the previous one. This consistent manifold is then passed onto a fixed decoding stage to produce robust output. The authors tested the method with two rhesus monkeys performing a cursor control task via an intracortical microelectrode array implanted in their primary motor cortex. Because neural instabilities occur unpredictably, the authors intentionally perturbed the system with known synthetic instabilities (such as adding firing-rate offsets, dropping electrodes and swapping activity between electrodes). Across 42 single-day experiments with new synthetic instabilities introduced each day, the authors observed that their stabilization method quickly restored the performance of the BCI (prior to stabilization, success rates had dropped by half and target-acquisition times increased at least fourfold) and within approximately 90 seconds, performance approached baseline levels. Without stabilization, the performance never returned to baseline levels within a session, indicating that the stabilizer addressed instabilities that the monkeys were unable to overcome on their own. The authors also show that the stabilizer counteracted both known synthetic and naturally occurring instabilities over 5-day experiments. However, the recovery from synthetic instabilities in these experiments took more than four times longer than that for the single-day sessions. This is partially due to the longer window of neural activity needed to estimate the manifold in multi-day experiments. Also, manifold stabilization yielded decoded command signals that were more accurate than those from an offline analysis using semi-supervised recalibration (which relies on knowing the monkey's intended target). This is in line with the fact that the intended target information accounts for just a small fraction of the activity of the neural population and shows that neural manifolds provide a richer signal[14].

Yu and co-authors' results are an online demonstration of robust recovery of BCI cursor control via manifold alignment (which had been shown previously in an offline setting[11,13]). Their study highlights the need for the further investigation of a number of translationally relevant questions, such as the degree to which the manifold structure is conserved across diverse naturalistic behaviours (in the authors' study, the monkeys performed a simple two-degrees-of-freedom cursor movement task); if other behaviours elicit a higher-dimensional manifold structure, its alignment might require the user to execute a fixed set of calibration behaviours. This in turn raises the question of how well the manifold alignment would generalize. Furthermore, although manifold approaches may yield stable decoding for well-learned behaviours, the learning of new behaviours may fundamentally alter the manifold[9,16]. Maintaining stable decoding in the face of a behaviour-dependent or learning-dependent manifold structure may be impossible without resorting to supervised methods. Solutions to this may include the identification of poor decoding performance via the detection of corrective movements, the monitoring of error-related neural activity or the integration of semi-supervised approaches[15] that learn gradual changes in the manifold-to-behaviour mapping. Moreover, a particular drawback of the authors' method is that the alignment of successive windows of neural activity requires that the activity recorded by a subset of electrodes remains stable across two successive windows, which may allow the manifold alignment and thus the decoder's performance to drift. If any given behaviour generates neural data spanning only a limited portion of the full manifold structure (which seems likely), it may be necessary to collect alignment data from many different behaviours within a time window that retains a sufficient number of stable electrodes. The fewer the available electrodes, the worse this problem becomes. Ultimately, an alignment approach that compares any current data distribution to a large reference distribution without requiring stable electrodes may prove advantageous.

Yu and co-authors used linear methods to robustly align successive windows of neural activity because they are computationally tractable for real-time applications. However, the dependence on successive-window alignment could be removed with alternative strategies. For example, a neural network that implemented adversarial domain adaptation achieved accurate manifold alignment and offline decoding with surprisingly small data windows and without requiring stable electrodes[13]. Also, approaches for dynamical systems that learn the rules that govern how neural activity evolves over time can provide complementary information to manifold structure that may further improve alignment[10,12]. Overall, further improvements to the stabilization of neural decoder performance will help mitigate the robustness issues that are hampering progress in the development of BCIs that can enable individuals with paralysis to regain function and independence. ❐

Lahiru N. Wimalasena[1], Lee E. Miller[2,3,4] and Chethan Pandarinath[1,5] ✉

*[1]Wallace H. Coulter Department of Biomedical Engineering, Emory University and Georgia Institute of Technology, Atlanta, GA, USA. [2]Physiology Department, Northwestern University, Chicago, IL, USA. [3]Department of Physical Medicine and Rehabilitation, Northwestern University, Chicago, IL, USA. [4]Department of Biomedical Engineering, Northwestern University, Evanston, IL, USA. [5]Emory Neuromodulation and Technology Innovation Center, Emory University, Atlanta, GA, USA.*

✉*e-mail: chethan@gatech.edu*

Published online: 13 July 2020
https://doi.org/10.1038/s41551-020-0587-9

References
1. Pandarinath, C. et al. *eLife* **6**, 18554 (2017).
2. Collinger, J. L. et al. *Lancet* **381**, 557–564 (2013).
3. Anumanchipalli, G. K., Chartier, J. & Chang, E. F. *Nature* **568**, 493–498 (2019).
4. Ajiboye, A. B. et al. *Lancet* **389**, 1821–1830 (2017).
5. Sussillo, D., Stavisky, S. D., Kao, J. C., Ryu, S. I. & Shenoy, K. V. *Nat. Commun.* **7**, 13749 (2016).
6. Downey, J. E., Schwed, N., Chase, S. M., Schwartz, A. B. & Collinger, J. L. *J. Neural Eng.* **15**, 046016 (2018).
7. Degenhart, A. D. et al. *Nat. Biomed. Eng.* https://doi.org/10.1038/s41551-020-0542-9 (2020).
8. Gallego, J. A., Perich, M. G., Miller, L. E. & Solla, S. A. *Neuron* **94**, 978–984 (2017).
9. Sadtler, P. T. et al. *Nature* **512**, 423–426 (2014).
10. Kao, J. C., Ryu, S. I. & Shenoy, K. V. *Sci. Rep.* **7**, 7395 (2017).
11. Dyer, E. L. et al. *Nat. Biomed. Eng.* **1**, 967–976 (2017).
12. Pandarinath, C. et al. *Nat. Meth.* **15**, 805–815 (2018).
13. Farshchian, A. et al. In *International Conference on Learning Representations* (OpenReview.net, 2019).
14. Gallego, J. A., Perich, M. G., Chowdhury, R. H., Solla, S. A. & Miller, L. E. *Nat. Neurosci.* **23**, 260–270 (2020).
15. Jarosiewicz, B. et al. *Sci. Transl. Med.* **7**, 313ra179 (2015).
16. Oby, E. R. et al. *Proc. Natl Acad. Sci. USA* **116**, 15210–15215 (2019).