UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Single-Neuron and Population Methods to Study the Circuit-Level Cortical Mechanisms of Multisensory Processing

Montelisciani, L.; Dijkema, E.; Olcese, U.

DOI
10.1007/978-1-0716-4208-5_1

Publication date
2025

Document Version
Final published version

Published in
Identification, Characterization, and Manipulation of Neuronal Ensembles

License
Article 25fa Dutch Copyright Act (https://www.openaccess.nl/en/policies/open-access-in-dutch-copyright-law-taverne-amendment)

Link to publication

Citation for published version (APA):
Montelisciani, L., Dijkema, E., & Olcese, U. (2025). Single-Neuron and Population Methods to Study the Circuit-Level Cortical Mechanisms of Multisensory Processing. In L. Carrillo-Reid (Ed.), *Identification, Characterization, and Manipulation of Neuronal Ensembles* (pp. 1-37). (Neuromethods). Humana Press. https://doi.org/10.1007/978-1-0716-4208-5_1

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.
UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:08 Nov 2026

Chapter 1

Single-Neuron and Population Methods to Study the Circuit-Level Cortical Mechanisms of Multisensory Processing

Luca Montelisciani, Eric Dijkema, and Umberto Olcese

Abstract

In the past decade, the study of the circuit-level mechanisms of cortical multisensory processing has made great progress, thanks to the development of methods to record and manipulate the activity of neuronal ensembles in head-fixed rodent models, in which the controlled delivery of sensory stimuli can be associated to behavioral tasks. However, compared to unisensory processing, the field is still in its infancy, and our understanding of how multiple sensory modalities are integrated and compared in the neocortex keeps being disrupted by novel findings. In this context, the emergence of tools to record and analyze the activity of large populations of neurons localized across multiple brain regions has emerged as a key strategy to link multivariate patterns of spiking activity to sensory stimuli and the associated behavioral responses. This chapter aims to provide a manual for researchers aiming to apply these techniques. Specifically, we provide an overview of the main experimental steps required to study multisensory processing in head-fixed mice with laminar electrophysiology. Furthermore, we discuss in depth the main preprocessing steps and the most relevant techniques used to analyze how single neurons and neuronal populations process sensory information.

Key words Multisensory processing, Electrophysiology, Ensemble activity, Population activity

1 Introduction

Our perception is intrinsically multisensory. When we drink our morning coffee, for example, we do not only taste it but also smell it, see it, and feel the warmth of the cup in our hands. Furthermore, flavor perception itself is also modulated by the other senses, for example, the color of what we eat or drink modifies the taste that we perceive [1, 2]. Many cognitive functions are also strongly influenced by multisensory factors. For instance, performance in a visual task is affected by auditory stimuli: a low-volume white noise enhances visual discrimination performance [3], while strong sounds typically impair visual detection [4]. Analogous effects are

Luis Carrillo-Reid (ed.), *Identification, Characterization, and Manipulation of Neuronal Ensembles*, Neuromethods, vol. 215,
https://doi.org/10.1007/978-1-0716-4208-5_1,
© The Author(s), under exclusive license to Springer Science+Business Media, LLC, part of Springer Nature 2025

observed for visual influences in auditory tasks [5]. In spite of the pervasiveness of multisensory processing, it is therefore surprising that the study of the underlying neuron-level mechanisms in the cortex is still in its infancy, compared to the study of unisensory processing [6]. In fact, most of our knowledge about the way in which neurons encode multisensory information historically either came from studies performed in the superior colliculus [7] or—for cortical processing—was largely limited to the study of cue integration [8, 9]. This is surprising for two reasons. First, in mammals, most cognitive functions influenced by—or based upon—multisensory factors reside in areas other than the superior colliculus [10]. Second, multisensory processing entails more than just cue integration and also includes—among other factors—the segregation and prioritization of sensory stimuli [6, 11–14].

Nevertheless, in the past decade, the study of the circuit-level mechanisms of cortical multisensory processing has made great advances thanks to (1) the development of strategies to train rodents, in particular mice, for various types of multisensory tasks [6, 15], and (2) the availability of methods to record and manipulate neuronal activity with unprecedented precision across areas [16]. Such tools, corroborated by transgenic animal models, include optogenetics, two-photon calcium imaging, and, crucially, high-channel-count silicon probes. This latter approach, in particular following the advent of Neuropixels probes [17], allows to record the activity of up to hundreds of neurons located in multiple areas, across both superficial and deep regions, and with a temporal resolution capable of resolving the shape of single action potentials. In the context of multisensory processing, these are essential features. First, multisensory processing necessarily involves multiple areas and, for what pertains to cortical regions, a complex interplay between different layers [4, 18, 19]. Second, neuron-level interactions between signals related to stimuli coming from different sensory modalities occur at timescales that preclude the use of techniques with slower temporal resolutions. For example, interactions between visual and auditory or tactile signals significantly change based on whether the stimulus onset asynchronicity of the different stimuli varies by only 50 ms [18, 20]; the major implications of the fast timescales involved in multisensory processing are exemplified by the renown double flash illusion [21]. Third, it has recently been shown that not all neural responses underlie multisensory processing. For example, auditory responses in the primary visual cortex are actually, for a large fraction, corollary discharges evoked by stereotypical, involuntary motor reactions to auditory stimuli [22, 23]. In order to disentangle these components, it is essential to have recording techniques with high temporal resolution, so that the various sources of neuronal activity—that can differ by only tens of milliseconds [23]—can be disambiguated. Finally, while multisensory processing in the superior colliculus has

traditionally been studied in terms of single-neuron responses [7], cortical multisensory processing (including cue integration) is a phenomenon thought to involve population-level dynamics [9, 24, 25], for whose investigation large-scale recordings are necessary.

These factors also apply to other fields of study (as is the case, for example, for what pertains the influence of spontaneous motor activity on the activity of cortical areas [26] or on visual processing and decision making [27, 28]) but are especially relevant for multisensory processing. In particular, several methods for the analysis of single-neuron and population-level activity have become very commonly applied to silicon probe recordings in the context of studying cortical multisensory processing, as they allow to both capture ongoing dynamics and disentangle distinct sources of activity. This chapter will therefore focus on providing an overview of such analytical methods. First, we will discuss the material and methodological requirements to study multisensory processing in head-fixed mice. This approach has become especially relevant because it allows to combine controlled delivery of sensory stimuli, monitoring of behavioral responses and internal state (e.g., arousal), and simultaneous application of multi-area recordings and optogenetics [13, 27]. A proper implementation of head-fixed recordings is essential to obtain high-quality data, as is the application of proper preprocessing techniques such as spike sorting. Spike sorting is essential to assign individual action potentials to single neurons, which is a prerequisite to study the neurophysiology of sensory processing. For this reason, a section of the chapter will be devoted to this topic. We will then focus on single-neuron methods for the analysis of sensory-evoked responses: classical peristimulus time histograms (PSTHs) and more advanced approaches to decode experimental variables from single-neuron activity or study neuronal encoding (e.g., generalized linear models). These approaches are essential to go beyond the level of firing rates and disentangle the various components embedded in spike trains. Finally, we will focus on population-level methods, for what pertains both dimensionality reduction approaches [29] as well as decoding methods. These techniques exploit the high-dimensional nature of large-scale recordings to uncover low-dimensional latent variables jointly encoded by neuronal ensembles [30]. Overall, this chapter will hopefully provide a useful resource for studying the circuit-level mechanisms of multisensory processing.

2 Materials and Experimental Paradigm

The following section is a step-by-step guide to perform an in vivo multisensory research paradigm in mice. It may function as a reference for those who wish to perform similar experiments or as a

template for in vivo systems neuroscience approaches, in general. Methods are described in chronological order of execution, encompassing all general steps from animal habituation to electrophysiological recordings. For applied examples of these methods, the reader is referred to Oude Lohuis et al. (2022) [13]. Due to their diverse nature and complexity, neuronal manipulations such as chemo- and optogenetics are beyond the scope of this chapter.

2.1 Habituation

Mice should be at least 8 weeks of age before starting any kind of behavioral or surgical procedure that does not concern developmental phenomena. This ensures that all sensory systems are mature and that any subsequent procedure is not affected by remaining body growth.

The first step of any experimental protocol should be habituation: animals need to get comfortable with being picked up and manipulated by experimenters. It is advisable to do so in a gradual manner, via which each animal is exposed to a gradually increasing extent of handling each day over a period of 1–2 weeks. Habituation can start with a researcher simply placing their hand in the home cage and letting mice get accustomed. Due to their exploratory nature, mice will approach the hand and sniff it, which is quickly followed by a retraction and return to their nest. Animals may be lured to approach by offering a sunflower seed or drop of milk. After 2–3 days, the handler will observe a change in behavior and animals will not hide from their hand anymore. The next step involves picking up the mice. Here, our preferred method is handling them with the use of a tube. Handling tubes are placed in each cage, so that animals will consider it a part of their environment and a safe place to hide. Often, tubes of red plastic are used, since they are translucent to the handler but not the animal, which sees the tube as a secluded and safe shelter. Once a mouse enters the tube, it may be removed from the cage. Simply lifting the tube to a distance where mice can still comfortably leave it (approximately 5 cm) to return to the cage is a good first step. Subsequent habituation steps include placing the tube on the arm of a handler, all the while keeping their hand in the cage so mice may traverse the arm to return to the home cage. Mice may be considered "habituated" once they can be placed on the arm or hands of a handler for a minute without showing signs of distress such as hyperactivity, aggression, or jumping.

2.2 Headbars for Head Restriction and Electrophysiology

In most (multi-)sensory research, controlling stimulus delivery is of prime importance [15]. This is achieved by head-fixation of subjects and requires the implantation of a so-called headbar. These headbars will almost always include a recording chamber—an opening that allows access to the skull without requiring any additional surgery. Headbars come in many shapes and sizes depending on the

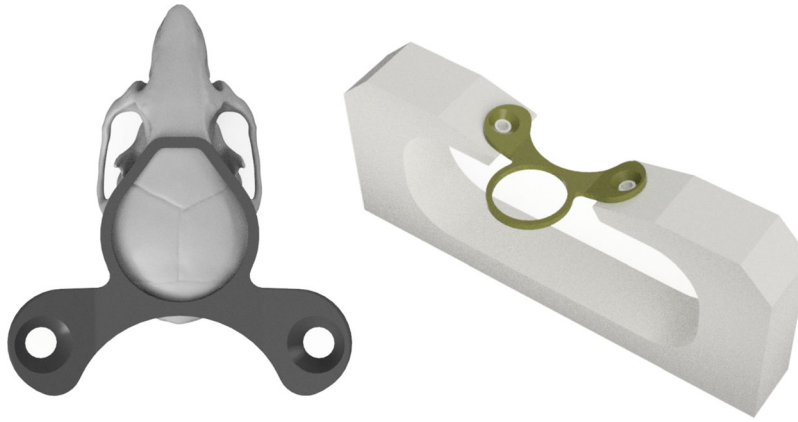

Fig. 1 Headbar example. Left: Rendering of a headbar superimposed on a mouse skull. This specific design allows simultaneous implantation of probes into posterior and anterior cortical areas. Right: Rendering of a similar headbar as it attaches to its complementary holder. "Ears" function as screw holes. Exposed skull is covered with a silicon elastomer and metal plate to prevent damage

animal species and experimental requirements such as area(s) of interest (see, e.g., Fig. 1).

Headbar surgery is performed under anesthesia and analgesia. A preoperative intraperitoneal injection of carprofen (10 mg/kg)—or another analgesic—is administered, before anesthesia is induced, typically with 3–5% of isoflurane. The animal is then moved to a nose clamp to ensure stability, where anesthesia is maintained with 1–1.5% isoflurane. Other anesthetics can of course be used. When deep anesthesia is achieved, confirmed by the absence of a pain reflex to a toe-pinch, an incision is made on the scalp using a tweezer and surgical scissors. Then, the periosteum is separated from the skull using a cotton swab. When a large headbar is implanted, it may be necessary to detach muscles such as pseudotemporalis profundus from the skull using a scalpel. When the skull's surface has been cleaned and dried, the headbar is attached to the skull using a glue (either regular super glue or glue for dental implants), after which a permanent protective layer is applied, usually in the form of dental cement. The bond between the skull and headbar will have to be able to withstand significant mechanical strain, so proper adhesion is crucial. Detachment of a headbar is considered a humane endpoint and thus requires the immediate sacrifice of the animal. It is therefore very important to ensure that the entire skull-implant interface is filled with cement. Following implantation, a small layer of cyanoacrylate glue is applied to the exposed skull, which is crucial to preserve the bone and prevent infection. The recording chamber is then sealed with a silicon elastomer and covered with a metal plate to avoid scratching. Subjects should be allowed to recover for at least 72 h after surgery, and ad libitum water, food pellets, and soft food should be available.

During this time, animals should be monitored for welfare issues, and analgesia should be administered appropriately. This can be done by adding carprofen to the water bottle, as it has been shown to keep steady blood levels of the drug without the need for invasive injections [31].

2.3 Dietary Restrictions for Task Performance

To motivate animals to respond to behavioral trials, some type of dietary restriction is necessary. For mice, water restriction is more commonly used than food restriction. By limiting the water intake of mice, both the total number of trials performed and the hit rate increase [32]. Water restriction should be introduced gradually, with the availability of hydration decreasing from ad libitum to approximately 1 mL per day for a typical adult mouse (with a minimum of 0,025 mL/g of body weight [33]) over the course of a week. At 1 mL per day, body weight is generally stable without animals showing signs of distress or discomfort. A convenient alternative to liquid water, whose administration is hard to control, is hydrogels. Hydrogels are materials that can contain up to 99% of water by volume, and gels for rodent consumption are commercially available. During the initial stages of water restriction. Hydrogel can be used to habituate animals to water restriction, supplement water during days with no training, or when animals did not collect their minimum daily hydration during training (approximately 1 mL). When animals are not being trained (i.e., during the weekend), an alternative to giving hydrogel is citric acid water. By adding 2–3% citric acid to water bottles, mice self-regulate their water intake without impacting behavioral performance [34].

2.4 Audiovisual Change Detection Task

There exists a rich variety of rodent behavioral tasks, ranging from freely moving to head-restricted, detection- to memory-based, spanning all sensory modalities. The scope of this section is confined to an audiovisual change detection task that we developed for mice [13, 14, 23], but the described principles will apply to many other tasks as well [6, 15]. Mice are trained on an audiovisual change detection task, where a change in one sensory domain is paired with licking a left lick spout and changes in the other domain to the right side. Before training, animals are habituated to head restriction in a gradual manner, increasing the duration of head restriction over the period of approximately 5 days. During the habituation phase, no stimuli are presented but licking the spout does result in the delivery of a small reward, familiarizing animals with collecting their daily hydration during head restriction. While restricted, the body of the animal is either in a plastic tube to align them to the screen or alternatively they can be placed on a floating Styrofoam ball to allow a certain degree of mobility. The latter approach is gaining popularity, especially in virtual reality experiments, where the movement of the ball is tracked and used to

navigate virtual environments [35–37]. After habituation, animals learn two separate unimodal change detection tasks before combining them into the multimodal task. By doing so, we strive to teach the task in small and easily learnable increments. Only when they perform well on the hardest version of a unimodal task do they progress to the next modality. The initial stages of training involve passive reward delivery paired with each stimulus change. When animals have learned to lick after a large fraction of the stimulus changes (~80%), typically after 2–3 days, rewards are only delivered following a correct lick. The d-prime (d'), a measure adapted from signal detection theory, is used as a metric for behavioral performance [38]. Animals progress between training stages when some criterium is met (e.g., d' > 0.5 for 3 consecutive sessions), in addition to exhibiting a sigmoid psychometric curve to confirm they are in fact responding to stimuli as a function of saliency (Fig. 2). Generally, mice will learn auditory and tactile tasks quickly and take longer to learn visual tasks.

Whenever a behavioral task involves multiple stimulus modalities and choices, possible biases toward specific modalities need to be accounted for. During multimodal training, a preference toward one of the stimulus sensory modalities may appear, resulting in the animal predominantly reacting to one modality. It is crucial to

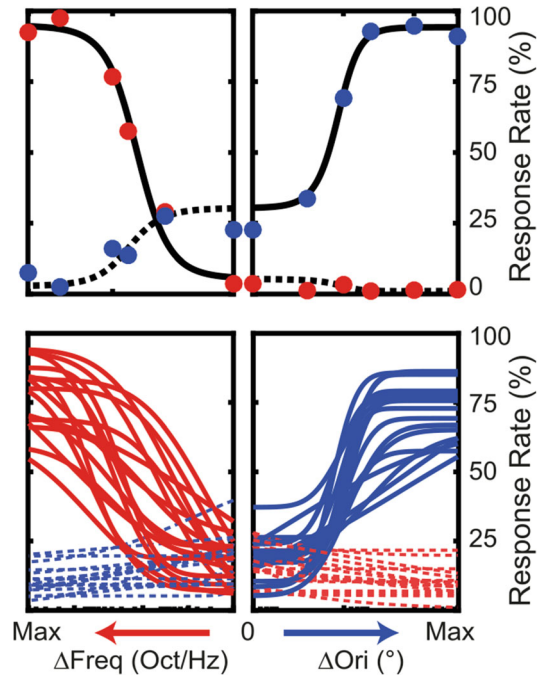

Fig. 2 Audiovisual change detection task psychometrics. Top panel shows response rates (dots) and model fit (line) for an example session. Bottom panel shows average psychometrics for multiple mice. (Adapted with permission from Ref. [13])

address such imbalances, for example, by changing stimulus saliency, lick detector distance, reward size, or the number of trials in a session for each modality. Additionally, individual subjects differ in their performance on each modality, so some type of normalization is necessary to compare between animals. We do so by determining a threshold stimulus, that is the stimulus at which 50% of trials are responded to correctly. During the experiment, each animal is presented with its respective threshold stimulus and a maximum saliency stimulus, e.g., a 90° orientation change of a drifting grating.

Analyses of behavioral data related to multisensory tasks also need to take biases into account. This is aided by the inclusion of probe trials into the task, i.e., trials where no stimulus is presented (Fig. 3). By comparing the animals' choices between stimulus- and probe trial, a better assessment of task performance can be made than simply looking at hit rates. The previously mentioned d' is a convenient measure of an observer's sensitivity to a signal but is limited to binary yes/no decisions in its original form. However, a modified version that allows the analyses multivariate decision tasks has been proposed [39].

2.5 Experimental Setup

This section will discuss several aspects related to experimental setups used in electrophysiological recordings in behaving rodents. Properly engineering and maintaining experimental setups is crucial to collecting high-quality, repeatable data. This is especially important because failure of a single device may result in the invalidation of the entire experimental session. For example, in the case of in vivo electrophysiology in behaving animals, failure to register the timestamps of stimuli on the same clock as the probe signal renders the data practically useless.

When it comes to choice detection, head-restricted rodent experiments often use lick detectors; animals lick to give their response to a trial, which is detected by a specialized device. Two common mechanisms used for lick detection are capacitive and

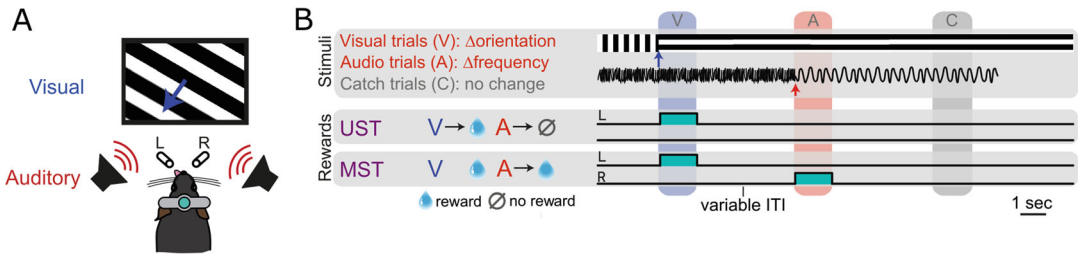

Fig. 3 Task setup. (**a**): Schematic overview of an audiovisual change detection task setup. Mice observe a drifting grating and auditory tone and are trained to detect changes in either one or both modalities. (**b**): Trial structure of visual, auditory, and catch trials. In visual trials, licks to, e.g., the left side are rewarded, whereas auditory trials are rewarded for right licks. During catch trials, no stimulus change is presented and animals are expected not to lick. Unisensory trained animals (UST) are only rewarded for one modality. (Adapted with permission from Ref. [13])

piezoelectric sensing. Capacitive detectors measure the capacitance of a conductive material, in this case, a lickspout. Whenever another conductive medium (e.g., a mouse) touches the spout, the capacitance changes and is registered as a lick. Capacitive lick detectors are robust and easy to use but not suitable for electrophysiology experiments, because each time the spout is licked, a tiny current runs through the animal which can be picked up as an electrical artifact. For this reason, sensors leveraging the piezo-electric effect are often used during experiments—lick spouts are attached to piezo-electric discs that generate a voltage whenever they are deformed. However, these devices are much more fragile and sensitive and thus inconvenient for daily training use.

A primary concern of all electrophysiologists is the elimination of electromagnetic interference from sources external to the brain. To isolate the subject from noise, experiments are usually performed inside a faraday cage. It is also of utmost importance to keep a proper ground tree, meaning all ground wires used in the setup lead to a single junction, instead of forming loops. Small currents can start to run in ground loops and introduce significant noise. Loaded power lines and AC-powered devices can cause noise several orders of magnitude larger than the signals of interest, which are typically in the micro- to millivolt range. These sources can easily be recognized by their 50/60 Hz waveform. After eliminating the 50/60 Hz noise component, residual noise sources may be isolated by carefully covering devices with a grounded material or placing them further away from the electrodes. Some devices, such as transformers and automatic switching devices (i.e., containing relays) can cause particularly large noise, and their use should be avoided if possible. In case digitization of neural signals does not happen at the headstage, it may also prove worthwhile to cover headstages in grounded aluminum foil and to use coaxed cables to connect to the acquisition system.

Another source of signal artifacts is animal motion. Motion artifacts are generated by a combination of electrical, magnetic, and ionic mechanisms [40], but all find their origin in the relative movement of the implanted probe with respect to the brain. For this type of noise, as with others, the best cure is prevention. Primarily, headbars should be securely attached to the holder so that body movement does not translate into head movement. Critical factors related to the probe are its leeway and the flexibility of the holder. Any movement of the probe may be disastrous to the integrity of the brain, as well as the probe itself. Ideally, both the animal platform and probe stand are attached to the same platform, for example, through magnetic clamps. Only very rigid materials should be used as a probe stand, so that touching probe stands doesn't result in movement. Relative movement between probes and the brain, so-called drift, may also be corrected post hoc during spike sorting. This method is discussed in more detail in this chapter's section on spike sorting.

2.6 Localization of Recording Sites: Mapping Retinotopy in Visual Cortex

A key aspect of multisensory research is to perform recordings in the correct brain regions to be studied. This, for large areas, can often be done via the use of stereotactic coordinates. However, when small regions need to be targeted or when subdivision of cortical areas (e.g., responding to a portion of the visual field) has to be reached, some form of functional mapping is required. For instance, to record cells that are tuned to visual stimuli, it is necessary to implant probes into the correct retinotopic location of V1. This can be guided by some type of retinotopic mapping, and in the context of this chapter, we will use this as a paradigmatic example of functional mapping. The most common options for retinotopic mapping are intrinsic optical signal (IOS, often also abbreviated as ISI) imaging or calcium imaging. Calcium imaging potentially provides a better spatiotemporal resolution and higher-quality mapping than IOS but requires the expression of genetically encoded calcium indicators, such as GCaMP6. IOS imaging utilizes the difference in red light absorption of blood as a function of oxygenation and is thus more widely applicable. For an extensive state-of-the-art retinotopic mapping method utilizing calcium imaging, the reader is referred to Zhuang et al. [41].

IOS imaging can be performed through the skull, without the use of an implanted cranial window. The procedure involves the careful thinning of the skull overlaying the area of interest, with the exception of animals young enough to image through the intact skull. The entire procedure is performed under anesthesia. When the bone is sufficiently thin, it becomes flexible to touch and translucent upon wetting. At this point, IOS imaging is possible. Stimuli equal to the experimental paradigm are presented, during which the cortex is illuminated by a red laser and frames are acquired by a camera. There are several factors that will determine the quality of the obtained signal:

- *Depth of anesthesia*—It is advisable to use low concentrations of anesthetics during imaging, e.g., 1% isoflurane. Deep anesthesia drastically reduces neural responses.
- *Smoothness of thinned skull*—If the thinned skull contains bumps or dents, light may scatter and introduce artifacts in the optical signal. Using larger drill bits can help to avoid this problem.
- *Body temperature*—It is pertinent to keep the animal at physiological body temperature, both for the well-being of the animal and the quality of the signal. A 1 °C deviation from homeostasis is enough to degrade the signal.
- *Camera focal point*—The camera must be focused on the tissue of interest. This is best achieved by using an animal platform with an adjustable z-axis. First, the platform is moved to a

position where the skull is in focus. Then, the platform should be moved up approximately 0.5–1 mm, depending on the thickness of the skull. This should place the focal plane of the camera within the cortex.

After acquisition, the hemodynamic response can be visualized by averaging each sequential frame over trials and then subtracting a no-stimulus baseline. A full description of an IOS experiment is beyond the scope of this chapter but is provided by Juavinett et al. [42]. When the area responsive to visual stimuli has been identified, a craniotomy is cut at that location using a scalpel. Animals should recover for at least 24 h before initiating awake experiments.

2.7 Laminar Electrophysiology

Once the area of interest has been identified, a common technique to measure in vivo neural activity is laminar-probe electrophysiology. Silicon probes can achieve very high spatial and temporal resolution recordings of both local field potential (LFP) and extracellular spiking activity. Especially with the advent of ultrahigh-density probes such as Neuropixels [43, 44] and SiNAPS [45] probes, hundreds of single units may be recorded simultaneously from both cortical and subcortical areas. In addition to single-shank probes, most manufacturers offer multi-shank alternatives for 2- or 3-dimensional recordings [44, 46, 47]. Such arrays or grids of electrodes offer the additional benefit of recording neurons at multiple closely spaced locations, which may offer insights into the mechanisms of intra-area neural processing.

Extracellular electrophysiology allows for an almost direct measurement of neural activity; the method does not rely on a secondary measure such as the blood-oxygen-level-dependent (BOLD) signal as is the case in, e.g., fMRI or functional ultrasound (fUS) imaging or calcium transients for calcium imaging. Transmembrane currents are directly measured as a potential difference, and thus temporal delays between neural events and recording are negligible. This is a unique benefit of electrophysiology, and no current methods offer comparable spatiotemporal performance [48].

The laminar structure of the mammalian cortex has striking anatomical and functional features, which can readily be explored with silicon probes. Electrode sites are generally spaced spanning vertical distance of microns to millimeters, enabling the concurrent positioning of electrodes in the cortex and underlying subcortical structures. Layer identities may be estimated post hoc using current-source density (CSD) and multiunit activity (MUA) analyses [49], where peak MUA activity is associated with cortical layer 5, and a range of CSD profiles predict several cortical layers. Besides physiological landmarks, exact probe trajectories can be determined postmortem by applying a fluorescent dye on probes during the final day of recording. One example of such a dye is DiI, a fluorescent lipophilic dye that diffuses into the lipid bilayers of the cell

where it stably remains fluorescent [14]. These combined methods allow researchers to precisely determine the position of electrodes (and consequently single neurons) within the cortical column and study the intracortical mechanisms of multisensory processing.

One limitation of probe electrophysiology is its invasiveness. Probes need to be implanted inside of the cortex through small craniotomies. As with every surgical procedure, extra care is required after surgery. During the experimental period, animal well-being should be intensely monitored. Probes can be implanted chronically [50] or acutely inserted and removed in subsequent recording sessions [13, 27]. In this case, each recording session will slightly damage the insertion area, especially with thicker probes. Multiple acute probe recordings can be done on different days, but inflammation and tissue damage will eventually forbid the continuation of experiments. These factors generally limit data collection to a period of approximately 5–10 days. Additionally, each acute recording session will contain spikes from a different set of neurons, as the probe is reinserted into at slightly different location. Some of the issues mentioned for acute probe recordings can be avoided by chronic implantation. Next generation probes (i.e., Neuropixels) have been successfully implanted for prolonged periods of time, in both rats and mice [50–52]. Chronic implantation eliminates the need for animals to be head-restricted during experiments, opening the door to experiments more closely resembling naturalistic behavior, tracking neurons over prolonged periods, whereas acute recording may only last for the duration of head restriction. Chronic implants have been shown to track the same neurons of days, allowing longitudinal studies to be performed at the single-neuron level [53].

Another drawback of using silicon probes in vivo is the limited available information about the recorded cells. There are currently no methods to univocally determine the identity of recorded neurons, such as in patch-clamp methods, where cells can be morphologically reconstructed. Most anatomical and physiological features of cells cannot readily be identified based purely on electrophysiological data, with the exception of their excitatory or inhibitory nature. A relatively common practice is the division of neurons based on their peak-to-trough delay, as inhibitory neurons have shorter waveforms than excitatory neurons [54–56]. Optogenetics can be used to elicit spiking activity from genetically identifiable neuronal populations, but this approach can be mostly applied to one neuronal subpopulation at a time [57].

Overall, the abovementioned procedures will enable to record ensemble activity (i.e., neuronal activity from multiple cells with single-neuron resolution) across the brain regions that, jointly, process activity coming from different sensory modalities.

3 Analysis of Neuronal Activity

After the previous description of experimental procedures, we now transition to methods for analyzing the neurophysiological data collected in the context of experiments meant to study multisensory processing. Specifically, we will focus on techniques enabling the study of topics such as the integration of visual and auditory cues and see how sound influences vision and vice versa [23], the influence of visual-tactile stimuli in decision-making tasks [58], and how multisensory stimuli are affected by animal behavior [22, 23]. Having collected a rich dataset that captures neural responses across different trials, modalities and time points, the next task is to distill this complexity into comprehensible and actionable insights. We will start with the preprocessing stage, specifically spike sorting. Next, we will discuss analysis methods at the level of single neurons. We will assess various methods such as raster plots, peri-stimulus time histograms (PSTH), tuning curves, receiver operating characteristic (ROC) curves, and generalized linear models (GLM). We will then conclude with population-level analyses, discussing the comparative advantages and disadvantages of two dimensionality reduction techniques and the use of population decoding techniques.

3.1 Preprocessing: Spike Sorting

Our initial focus in preprocessing the data is on spiking sorting, a set of procedures meant to isolate the activity of putative individual neuronal units, and thus enable the investigation of their responses to sensory stimuli [59]. This level of resolution is indispensable for elucidating the neurophysiological mechanisms that govern sensory encoding and multisensory integration. Today, various spike sorting algorithms are at our disposal, including but not limited to Kilosort [59, 60], Klusta [61], SpyKING [62], and IronClust [63]. All these methods can be used and compared via the use of tools such as SpikeInterface [64], a Python package designed to streamline the process of spike sorting and electrophysiological data analysis.

Before being able to apply spike sorting, the raw data needs to be preprocessed to make the detection of action potentials easier. Two common preprocessing steps are common median reference and band-pass filtering [17]. Common median (or mean) reference involves subtracting the median (mean) voltage of all active channels from each individual channel's raw voltage data, which helps reduce common noise sources and improves signal-to-noise ratio. Next, band-pass filtering is used to remove noise and other unwanted signal components (e.g., LFP). A typical frequency range is between 600 and 6000 Hz, although variations exist.

Delving into spike sorting, there are two main approaches that are currently used: waveform-based and template-based methods.

In waveform-based methods, the first step typically involves identifying spike events, often determined by voltage thresholds. Next, their waveforms are aligned to ensure consistent representation. These waveforms are then transformed into a shared low-dimensional space, often through techniques like principal component analysis (PCA) [65]. This feature extraction step simplifies the data while preserving essential features, setting the stage for clustering. Algorithms in this category then usually employ clustering techniques such as Gaussian mixtures or peak-density detection to group similar waveforms into putative single neurons [66].

In contrast, template-based methods such as Kilosort [60] take a different approach. These methods focus on initially generating a set of templates, which serve as typical spike shapes against which detected spikes can be compared. After an initial round of spike detection, the detected spikes are matched against these templates. This template matching is particularly effective in resolving overlapping spikes, where multiple spikes occur almost simultaneously. Finally, the spikes are grouped into distinct clusters, each representing an individual neuron, based on their similarity to the templates.

Regardless of the chosen approach, automated spike sorting has its limitations. It can mistakenly detect noise as spikes, get confused by overlapping spikes from multiple neurons, and misclassify signals due to small shifts in the electrode's position during the experiment, known as "probe drift." For these reasons a manual curation step is necessary after the initial automated phase. This serves three key purposes: first, it combines spike groups that the automated process incorrectly separated, assuming they originated from different neurons; second, it separates spike groups that the automated process wrongly identified as originating from just one neuron; finally, it filters out artefactual and mixed signals, often referred to as "multiunit activity," which are actually undistinguishable action potentials from multiple neurons (Fig. 4). Software tools—such as Phy [67]—have been developed to facilitate this process. During manual curation, operators initiate the process by scrutinizing neurons' spiking autocorrelogram, specifically focusing on a 2-ms window within the refractory period—the time following an action potential when the neuron cannot fire again—to ensure it is free of contamination. If there are no spurious spikes in the refractory period, one can then select the most similar cluster for comparison. Next, waveform shapes, amplitudes, and interspike interval distributions are examined and compared to spot cluster similarities. If two clusters appear very similar and originate from neighboring (or overlapping) recording channels, the clusters are merged. Finally, if the merged cluster shows consistent correlogram patterns, waveforms, amplitudes, and interspike intervals, it is classified as a "good" single unit; otherwise, it is labeled as a "multiunit" activity. Manual curation becomes particularly important

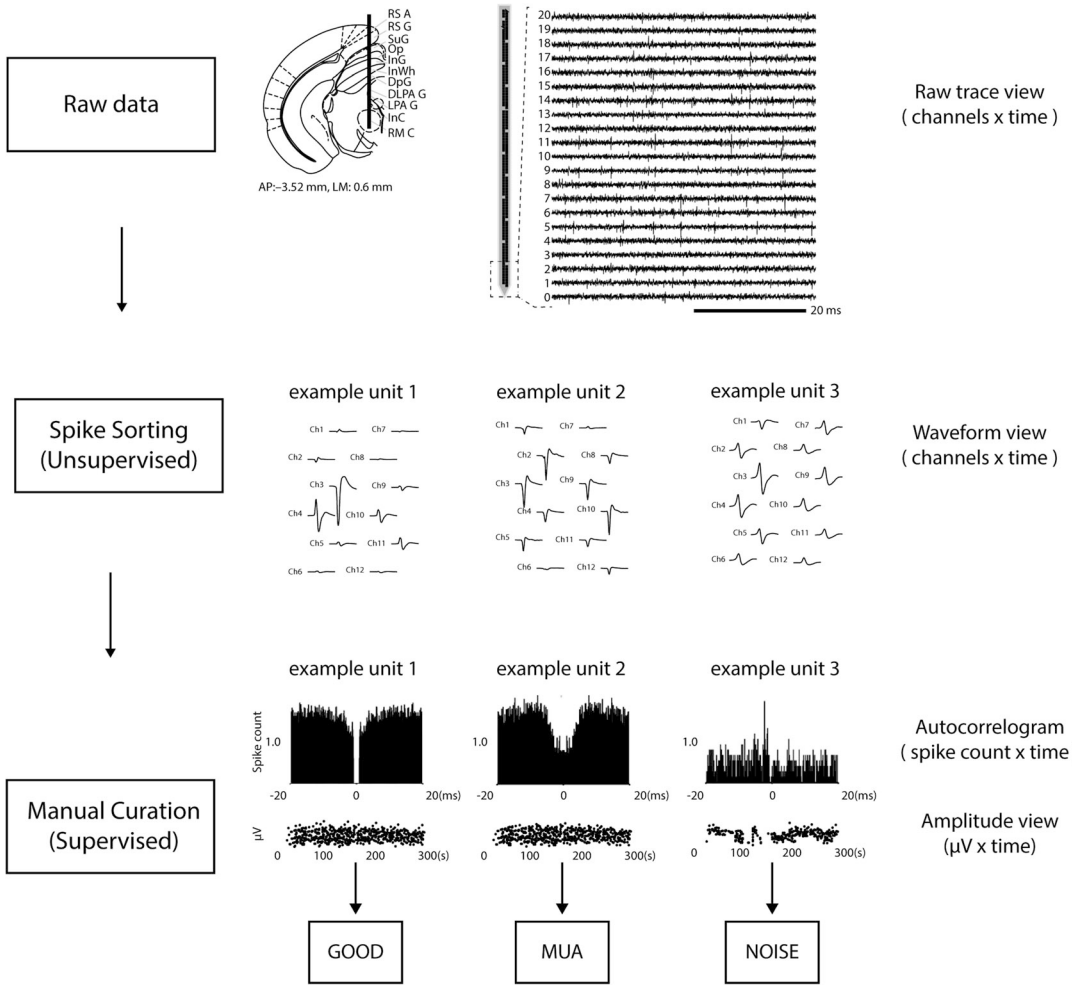

Fig. 4 Spike sorting pipeline: from data collection to manual curation. *Raw data*: This panel provides a schematic representation of the probe track reconstruction within the mouse brain. The targeted areas were situated at −3.52 mm anterior–posterior (AP) and 0.5 mm lateral–medial (LM) axis relative to bregma. Displayed on the right side are the filtered voltage traces, which have undergone a 300 Hz high-pass and 9000 Hz low-pass filter, spanning 50 ms from 20 sample channels located at the probe's tip. *Unsupervised spike sorting*: Here, the output for three example units is showcased. This allows for the visualization of unit activity at each channel by plotting the recorded waveforms. *Manual curation:* This final step involves a detailed examination of the three example neurons from the previous section. Using patterns observed in the autocorrelogram and amplitude, these neurons can be categorized as follows: "good" indicating single unit activity, "MUA" representing multiunit activity, or "noise." (Adapted with permission from Ref. [52])

when dealing with recordings presenting electrode drift. In these scenarios, the waveforms produced by a specific neuron change over time and could be incorrectly allocated to multiple clusters. While new preprocessing software, like Kilosort 3, performs drift corrections, older software lacks this operation [60]. Thus, the curation step helps to ensure these waveforms are correctly associated with their originating neuron.

While spike sorting algorithms and manual curation prepare the data for analysis, it is crucial to understand that these techniques have their own set of limitations.

Data originating from outdated laminar probes (with a low density of recording sites) or tetrodes are more susceptible to issues related to noise or drifting of signals over time. While modern software like SpikeInterface helps select the most appropriate algorithm for each probe, spike sorting algorithms still need to improve in fully addressing these challenges. Moreover, while often effective, manual curation can be both subjective and time-consuming, and its feasibility may be questioned when dealing with vast datasets, for instance, with data coming from neuropixels probes.

In conclusion, spike sorting is a pivotal step in the data analysis pipeline for neural recordings. Carefully processing and curating the data can provide a more precise and accurate picture of the neural activity underlying behavior and cognition. Although limitations exist, ongoing developments in algorithms and analysis methods are helping to address these challenges and further improve the quality and reliability of neural data.

3.2 Analysis of Single Neuron Evoked Responses

The brain's process of combining information from various sensory modalities to fully understand our environment relies on neural encoding and decoding mechanisms.

Encoding allows neurons to transform real-world sensory information from multiple sources into a form the brain can comprehend and use. Decoding, instead, is the process of interpreting these signals, enabling the brain to perceive a coherent, multisensory world representation based solely on the activity of neurons [68, 69].

To understand these processes, analyzing single-neuron activity is often the first required. This is in line with the Sherringtonian view that individual neurons serve as the foundational units for neural function, each contributing uniquely to the brain's overall ability to process and integrate information from multiple sensory modalities—but see [70].

To this end, several techniques for the visualization and quantification of neuronal responses come into play, facilitating both analysis and visualization. Among these techniques are raster plots, which display the temporal patterns of spikes in response to stimuli. Peri-stimulus time histograms (PSTHs) offer a quantified representation of neural responses over time intervals. Tuning curves map a neuron's firing rate in response to varying stimulus parameters. Receiver operating characteristic (ROC) curves are used to assess the discrimination ability of the neurons for different stimuli. Finally, generalized linear models (GLMs) allow for the statistical analysis of relationships between neural activity and external factors like stimuli or behavior. All these methods collectively provide a comprehensive framework for investigating the single-neuron underpinnings of information encoding and processing.

3.2.1 Raster Plots

A raster plot is arguably the simplest and most intuitive method to examine a single neuron's firing rate over time, allowing for visual inspection of the trial-by-trial variability of the time-locked responses to sensory stimuli.

Constructing a raster plot begins by capturing when spikes, or action potentials, occur. Next, these spike times are aligned to the times of particular events of interest, such as a sensory stimulus being presented. Once this alignment process is complete, time is plotted on the x-axis, and the trial number is on the y-axis. Each action potential is represented with a small vertical line or a dot at the corresponding time it occurred (Fig. 5a). As a result, the raster plot provides a clear, intuitive depiction of a neuron's activity across multiple trials relative to a given event, such as the presentation of a sensory stimulus. This visual representation allows researchers to easily identify patterns, response latencies, and variabilities in neuronal responses.

3.2.2 Peristimulus Time Histograms (PSTHs)

PSTHs are commonly used to quantify how a neuron's activity changes over time in response to a specific event or stimulus. This involves counting spikes in defined time intervals before, during, and after the stimulus (or event) onset. The PSTH's foundation is the raster plot, from which a PSTH is basically obtained by aggregating spikes within specific time bins across trial. This provides an average temporal profile of neuronal activity around the stimulus.

Drawing from the raster plot to construct a PSTH, the first step involves selecting an appropriate bin size. This decision is often based on the balance between resolution and statistical reliability. Large bins, like 50–100 ms, blur finer temporal patterns, while excessively narrow bins, like 1–5 ms, are intrinsically noisy. The choice also hinges on the distinction between rate and temporal codes. Broader bins might be sufficient for capturing overall firing rates without detailing precise spike timing. However, if discerning the exact timing of spikes is paramount, smaller bins in the range of 20–50 ms are preferable.

The spike data is then binned. Nonoverlapping binning divides the time range into equally sized sections. Instead, the sliding window method shifts a window incrementally across the time range, frequently overlapping with the previous window to capture subtler temporal nuances.

Once the data has been binned, the next step is smoothing. This process helps reduce noise and accentuate underlying patterns. Common smoothing techniques include Gaussian smoothing and moving average filters. However, it's important to balance smoothing with the preservation of key temporal details, like precise spike timing, that are critical for accurate data interpretation. Afterwards, normalization to spikes per second is always needed so that the PSTH can be used for comparisons across neurons, conditions, and experiments. Finally, z-scoring is often applied as it provides a more

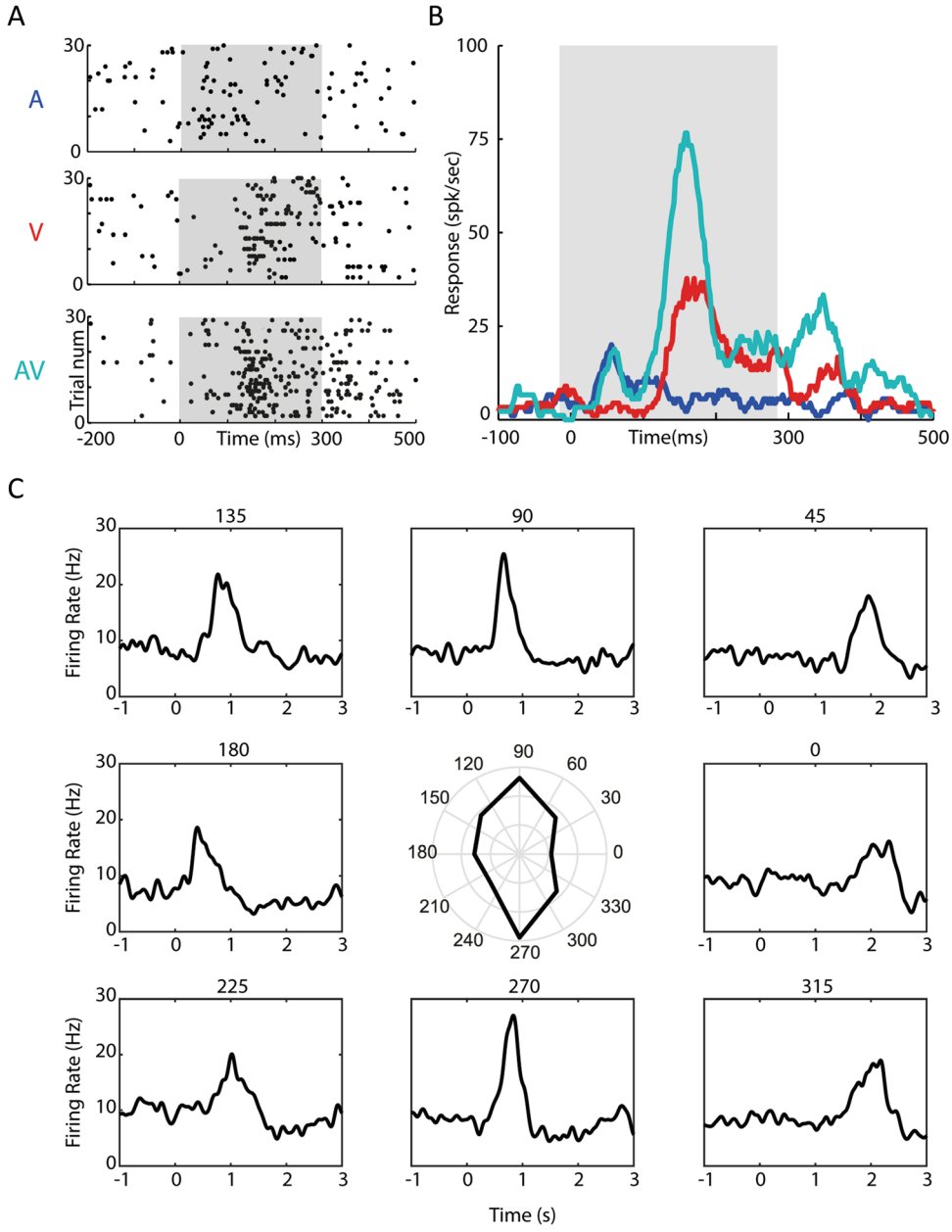

Fig. 5 Raster plot, PSTH, and orientation tuning after audiovisual stimulus presentation: (**a**): Raster plots of auditory (A, blue) visual (V, red) and audiovisual (AV, green) trials of an example neuron exhibiting multisensory enhancement. (**b**): PSTH of the same neuron shown in panel A demonstrating enhanced firing in response to audiovisual stimuli compared with unimodal stimuli (same colors as in panel A). (**c**): PSTH and polar plot demonstrating the orientation tuning and magnitude of response (Hz) of example neuron in V1 of mouse. (Adapted with permission from Refs. [71, 72])

standardized comparison of activity levels throughout the analyzed time window and neurons (Fig. 5b).

In multisensory analysis, PSTHs are useful tools not only to visualize but also to quantify specific features of neuronal activity. Metrics such as peak firing rate and latency to peak can be used to understand how neurons integrate inputs from multiple sensory modalities. For instance, a higher peak firing rate when auditory and visual stimuli are presented together compared to unisensory presentations indicates multisensory enhancement. On the other hand, a decreased peak firing rate indicates response depression. Besides, calculating the area under the PSTH curve provides the neuron's total activity in response to multisensory stimuli, further aiding in identifying enhancement or depression effects [73]. By thoroughly interpreting and quantifying these patterns, PSTHs evolve from mere visual representations to tools that enable preliminary measurements. These insights shed light on neuronal responses' timing, consistency, and magnitude, showing how neurons encode and differentiate information.

3.2.3 Tuning Curves

While PSTHs provide insights into a neuron's response to a single stimulus, tuning curves offer a broader perspective by comparing neuronal responses across multiple stimuli. In constructing a tuning curve, researchers present a series of stimuli during the experiment, such as gratings at different orientations (e.g., 0 to 180 degrees) or pure tones at varying frequencies (e.g., 5–80 kHz) [74]. A neuron is then considered "tuned" when the average response against varying feature values of the same stimulus class is significantly higher than others.

To analyze and construct the tuning curve analysis, a PSTH is first computed for each stimulus. Metrics such as mean firing rate or peak response are then calculated, separately for each stimulus/condition: the former by averaging spikes per bin for each stimulus and the latter by identifying the highest spike count usually during the stimulus presentation window. Polar plots or similar visualizations are often used to display how response magnitudes vary as a function of stimulus features (Fig. 5c). Successively, the stimulus value that induces the peak response from this plot is deemed the neuron's preferred stimulus, indicating the characteristic the neuron is most receptive to. Researchers also often fit a mathematical function, e.g., a Gaussian, cosine, or von Mises curve, to the observed data points. Such curve fitting enhances the estimation of a neuron's tuning but also equips scientists with the ability to predict neuronal responses to untested stimulus values within the tested range. Finally, statistical analyses are indispensable to validate that the identified tuning traits aren't merely random variances. Given the multiple stimuli conditions, an ANOVA is often employed to compare the neuron's response across different

orientations or frequencies against its baseline firing rate. Circular statistics are often employed in the visual domain.

Even though PSTHs and tuning curves provide valuable insights into neural activity, they come with limitations. For instance, they fail to capture the variability in neural responses across different trials or within a single trial. This limitation is particularly important in multisensory research, as it can mask interactions between different sensory inputs that may only become apparent in individual trials. Furthermore, these methods can miss the mixed effects in experimental data. For instance, movement artifacts have been observed to alter neural recordings when an animal moves during sound presentation [22, 23]. Physiological changes such as variations in heart rate or pupil size can also introduce noise into neural activity. Because PSTHs and tuning curves rely on averaged measures, these factors can be masked in the aggregation across trials, leading to potentially inaccurate or incomplete interpretations of neural behavior. For these reasons, techniques such as receiver operating characteristics (ROC) curves and generalized linear models (GLM) are often used to supplement raster plots, PSTHs, and tuning curves.

3.2.4 Receiver Operating Characteristic (ROC)

Receiver operating characteristic (ROC) analysis is a powerful tool for studying neuronal responses, particularly during multisensory integration experiments. ROC analysis can quantify a neuron's discriminative capacity, specifically its ability to classify or distinguish between sensory inputs correctly [13]. For instance, considering a neuron responding to both visual and auditory stimuli, while a PSTH might show the neuron's average firing rate in response to each stimulus, ROC analysis would assess at which point in time the neuron differentiates between the two stimuli [13].

ROC analysis requires several stages. After data collection, researchers gather raw data, focusing on the neurons' responses during each trial. Such spike trains are converted into continuous traces by computing single-trial PSTHs. After, setting the decision threshold is crucial. The "threshold" is a cutoff value used for classifying instances—based on their predicted probabilities or scores—as either positive or negative. Initially, this threshold is set at the maximum predicted score, meaning all instances are classified as negative. ROC analysis then proceeds by varying this threshold. By doing so, one can assess a model's performance at different levels of sensitivity and specificity. Sensitivity measures the rate of correctly identified positive instances while specificity measures the rate of correctly identified negative instances. Adjusting the threshold helps understand the trade-offs between these two metrics, optimizing a model's overall performance. After setting the various decision thresholds, the next step is to calculate key metrics at each

threshold level: the true positive rate (TPR) and the false positive rate (FPR). The TPR, also known as sensitivity, is the proportion of actual positive cases that are correctly identified by the model. On the other hand, the FPR is the proportion of actual negative cases that are incorrectly classified as positive by the model. With these rates, the visual representation of the ROC analysis, the ROC curve, takes form. In the graph, the FPR is on the x-axis, and the TPR is on the y-axis. Each coordinate on this curve represents a pair of FPR and TPR values associated with distinct thresholds. The ideal ROC curve tends to lean toward the upper-left corner of the graph, symbolizing a high degree of sensitivity and a minuscule rate of false positives (Fig. 6a).

The ROC curve is summarized into a singular metric: the area under the curve (AUC). An AUC of 1 represents perfect discrimination, while 0.5 indicates random guessing. To assess the significance of the AUC against random guessing, corresponding p-values are calculated [13]. A low p-value (typically below 0.05) would imply that the observed AUC is statistically significant and not due to random chance (Fig. 6a).

To visualize the time-resolved discriminative ability of neurons, one can plot the fraction of neurons significantly discriminating (encoding) between two conditions over time. For example, we can observe the time at which neurons in the V1 area encode

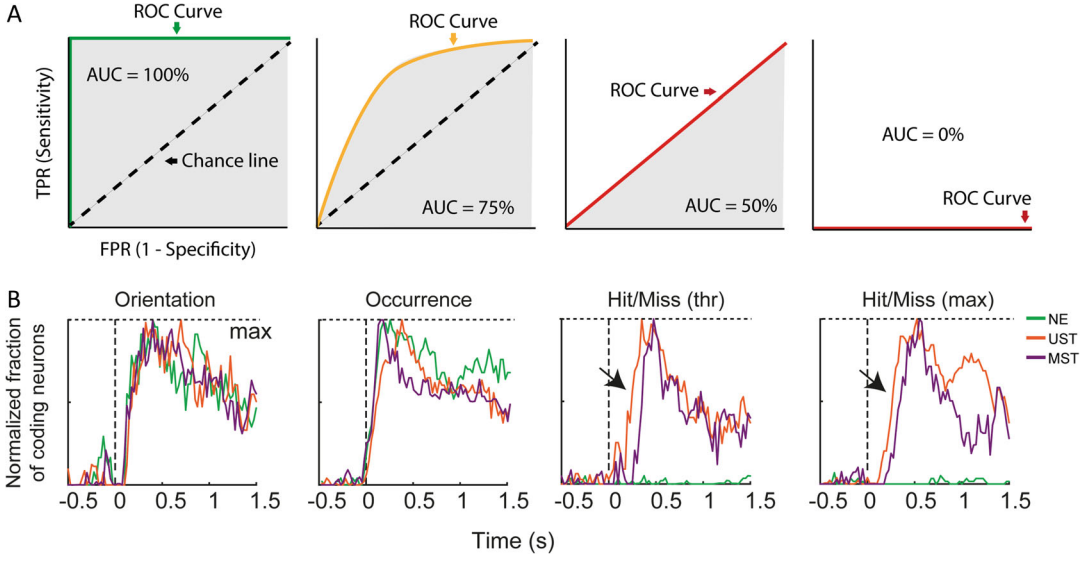

Fig. 6 ROC curve schematic and results: (**a**): ROC curves and area under the curve (AUC) for different discrimination levels. (**b**): We used a single-neuron decoding analysis (ROC) to compute the fraction of neurons (summed over all recordings) coding for task-relevant variables over time. The coding of change occurrence and orientation (visual features) showed similar dynamics across cohorts, whereas hit/miss coding (visual hits vs misses) was only present in UST and MST mice and started earlier in UST than MST mice (highlighted with black arrows). (Adapted with permission from Ref. [13])

different stimulus and task features during an audiovisual change detection task (Fig. 6b). Through ROC analysis, we can pinpoint the specific time points where neurons differentiate between "hit" and "miss" trials, thereby revealing when they are encoding task-relevant feature (Fig. 6b).

ROC analysis thus provides a quantifiable metric for the efficiency of neural encoding across different sensory inputs or conditions. However, ROC analysis has some limitations.

First, its design primarily suits binary classifiers, evaluating two distinct classes such as "response" and "no response." In the multisensory integration context, however, stimuli responses can span more than two categories [14].

Additionally, ROC assumes an equal distribution between the two classes it evaluates. Since neural datasets sometimes conform to this assumption, the AUC might paint an overly positive picture in cases lacking this balance. A skewed dataset might lead to biased classification, falsely elevating perceived accuracy [75]. However, this limitation can be addressed by subsampling one of the two classes to match the size of the other, thereby reducing the potential for biased outcomes.

3.2.5 Generalized Linear Models (GLMs)

While ROC analysis is used for decoding, primarily to assess how well a neuron can distinguish between two conditions, GLMs are encoding models that explore how sensory input, task conditions, and behavioral responses shape neuronal spiking activity [13, 23].

Essentially, a GLM links a dependent variable, such as a neuron's spiking activity, to various independent predictors like trial type, pupil size, or motion energy (here meant as pixel intensity difference between consecutive frames of a video's recording of orofacial movements [22, 23, 26].

Indeed, recent works showed how, in the context of multisensory experiments, it is essential to quantify how much of the observed neural variability is accounted for by the animal's movement, thereby isolating the neural responses that are truly attributable to the sensory inputs [22, 23]. This makes GLMs very useful for multifactorial neuronal activity analysis.

GLMs rely on probability distributions to capture the relationships in neural data. A probability distribution helps to describe how the outcome variable, such as spike count, is distributed across different conditions or time points. By selecting appropriate probability distributions, GLMs can be tailored to various research objectives and data types. These distributions can be Gaussian for continuous data, binomial for binary data, or Poisson for count data like spike-train data. Specifically, when working with spike-train data the Poisson GLM is often preferred, as it accurately models expected spike counts within short time intervals.

After choosing the model type, it is fundamental to choose an appropriate link function. This transformative element bridges nonlinear associations between the dependent variable and predictors into a linearizable format. The specific choice of a link function is informed by the nature of the data and the relationship under study. In the context of Poisson GLMs for spike-train data, the log link function is typically used, as it models the relationship between the expected log count (or rate) and the predictors:

$$\log \lambda_i = \beta_0 + \beta x_i$$

Here:

- $\log \lambda_i$ is the natural logarithm of the expected count or rate for the ith observation.
- β_0 is the intercept, representing the expected log count or rate when the predictor is zero.
- β is the coefficient quantifying the change in the expected log count or rate for a one-unit increase in the predictor.
- x_i is the value of the predictor for the ith observation.

After selecting the appropriate link function, the next crucial step is to choose the right predictors, which are variables believed to influence neural responses. These predictors can vary widely depending on the research context. For example, in audiovisual experiments aimed at studying multisensory integration, potential predictors could include the timing and intensity of a visual stimulus, the frequency of an auditory stimulus, and behavioral metrics such as eye movements. The model includes these predictors to help account for the variability in neural firing rates and better understand the underlying neural mechanisms (Fig. 7b). Next, single-trial PSTHs for each unit across all trials are generated as this data serves as the foundation for fitting the full model.

The model is then fitted to the data through a process that includes cross-validation. Specifically, cross-validation is carried out over individual trials to fine-tune model parameters and ensure the model's predictive accuracy and robustness. Once the optimal model is identified, a shuffling process is applied to the predictors across various categories. This is done to account for any spurious relationships in the data. The validity of the shuffling is estimated by comparing the model's predictive performance before and after shuffling, helping to distinguish the relationships from those that may have arisen by chance.

Upon completing the analysis, it is important to visually represent the results to discern how our predictors elucidate the variances in firing rates. Typically, plotting how firing rates are explained by each predictor gives a precise view of neural activity patterns and their temporal relationship with the predictors (Fig. 7d). Bar plots can be used for visualization of the contribution

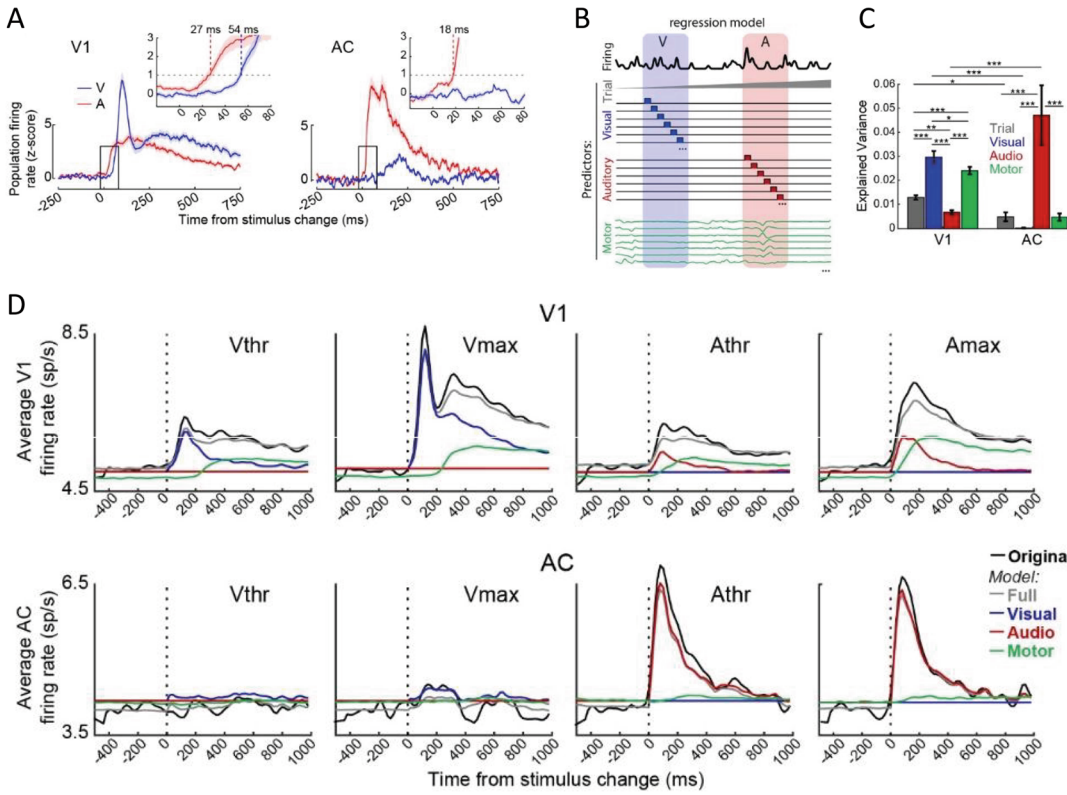

Fig. 7 GLM schematic and output from silicon probe recording in V1 and AC mouse during audiovisual stimulation: (**a**): Population firing rate after auditory and visual stimulus changes in V1 (left) and AC (right). (**b**): Schematic of the model. Single-neuron firing rate was predicted as a linear combination of four sets of predictors. A "trial" predictor captured rate fluctuation throughout the session. For visual (V) and auditory (A) stimuli, a set of separate predictors spanned 1 s after stimulus change per saliency level and feature (orientation and auditory frequency; shown only for one trial type) to capture neuron-specific response patterns. Finally, motor predictors were also included. (**c**): Explained variance during 0–200 ms of all trials for each category of predictors. Mean ± SEM across neurons. (**d**): The firing rate of single neurons was predicted using only subsets of predictors (visual, blue; auditory, red; motor, green) or all predictors (gray) and compared to the original firing rate (black). Shown is the firing rate during different trial types averaged across V1 neurons (upper row) or AC neurons (bottom row). (Adapted with permission from Ref. [23])

of individual predictors (Fig. 7c). Such visual aids not only facilitate a comprehensive understanding of the model's outputs but also provide a rigorous validation of the findings in a format that's both accessible and academically rigorous.

However, GLMs also have limitations. One key limitation is the assumption of a linear relationship between predictors, e.g., the intensity of sensory stimuli, and the response variable, usually neuronal activity. This assumption may not hold in multisensory settings where stimuli from different senses interact complexly to influence neuronal activity. If an incorrect link function is chosen, the model might offer misleading insights into how sensory inputs

affect neuronal responses, potentially resulting in flawed conclusions or misdirected future research.

Furthermore, creating the right set of predictors can be challenging. In a multisensory context, these predictors might include factors such as individual trials, variations in stimuli, and related movements. Constructing an appropriate design matrix is crucial: missing important predictors can hide significant effects and reduce sensitivity, while adding unnecessary ones can introduce noise [76].

Lastly, GLMs are sensitive to outliers. Specifically, rare but extreme neuronal responses can disproportionately influence the model. If not addressed, this can make the model less generalizable, resulting in erroneous interpretations of what constitutes "typical" neuronal behavior in response to multisensory stimuli. While robust regression techniques can mitigate this issue, they also introduce additional complexities to the model's architecture, which might complicate its interpretation and application [76].

To conclude this section on the analysis of single-neuron activity, studying the firing properties of individual neurons is crucial for analyzing neural activity, particularly in multisensory experiments. While traditional methods like raster plots, PSTHs, and tuning curves offer basic understanding—they help elucidate the general patterns and primary characteristics of neural activity—they often simplify what neural responses actually reflect. These methods might not fully capture between trial dynamics and the interaction effects between stimuli from different sensory modalities. ROC analysis adds valuable insight by quantifying a neuron's ability to distinguish between different sensory stimuli, excelling in measuring the sensitivity of neuronal responses. However, in multisensory experiments, ROC may face challenges incorporating behavioral aspects like animal movement or licking. For instance, neural activity recorded during a multisensory trial may not solely be influenced by the sensory stimuli presented but also by the animal's behavior at that time [77].

On the other hand, GLMs can elucidate the relationship between a neuron's activity and multiple factors, such as trial type and behavior, making them well-suited for multisensory experiments where multiple inputs interact to influence neuronal responses. However, GLMs operate under certain assumptions, such as a linear relationship between predictors and response and independently distributed errors, which might only sometimes hold since they can also be sensitive to outliers, leading to poorly generalized models.

3.3 Analysis of Population Evoked Responses

After having discussed single-neuron analysis methods, we will now examine population-level dynamics. A philosophical divide distinguishes single-neuron analysis and population-level analysis. The former, rooted in reductionism, focuses on single neurons as the basic computational units of the brain. In contrast, population-level

analysis embraces holism, focusing on the collective dynamics of interconnected neural networks and emergent properties [30, 70, 78]. While single-neuron research sheds light on cell-specific activity, population studies offer a broader perspective on how groups of neurons collectively process information. As we categorize the techniques employed in the latter analysis, two primary methodologies emerge, based on, respectively, unsupervised and supervised learning.

Unsupervised strategies, such as principal component analysis (PCA) and tensor component analysis (TCA), are typically dimensionality reduction techniques, and they explore and reveal latent variables within multisensory data, free from the boundaries of predetermined categories. In contrast, supervised techniques, particularly neural decoding, quantify how much information about sensory or behavioral events is embedded within neural firing patterns [14].

In an experiment in which animals must choose between two choices to obtain a reward, neurons may encode multiple aspects of the task, such as the choices presented, the animal's decision, and its confidence in that decision [79]. This results in a "mixed" representation of different task elements by the same set of neurons, making the data challenging to analyze. Dimensionality reduction techniques are instrumental here, systematically capturing most of the data's variance and elucidating how neural representations depend on various task parameters like stimuli, decisions, or rewards. Decoding techniques instead infer variables such as stimuli or behavioral states from observed neural activity. By training models on specific neural patterns aligned with known stimuli or behaviors, decoding can predict these factors from neural datasets. For instance, decoding neural activity in the visual cortex allows to quantify how well neuronal activity in this area can be used to predict the type of visual stimulus presented to an organism [71]. Similarly, decoding techniques have revolutionized brain-machine interface applications when applied to the motor cortex [80], facilitating direct communication between neural circuits and external machinery based on interpreted neural representations of intended actions.

3.3.1 Dimensionality Reduction Techniques

Principal component analysis (PCA) and tensor component analysis (TCA) stand out as techniques employed in neuroscience for dimensionality reduction and extraction of latent variables. PCA is probably the most commonly used tool for reducing the dimensionality of neural data, transforming it into a set of orthogonal axes that capture the most significant patterns. For instance, in a decision-making task with audiovisual stimuli, PCA not only de-noises the data but also isolates dimensions that capture the variance specific to movement and decisions [81]. TCA can be seen as a multidimensional extension of PCA [82]. For example,

TCA can uncover latent dimensions corresponding to multiple task-relevant components (e.g., sensory and motor components), operating simultaneously across neurons, trials, and multiple trial types [14].

3.3.2 Principal Component Analysis (PCA)

Principal component analysis (PCA) is a statistical method that identifies and generates the primary axes or dimensions along which the data varies the most. By doing so, it creates new variables, called principal components, which are linear combinations of the original features. These principal components effectively simplify the dataset's complexity while retaining as much of the original variance as possible. Besides, the two components are orthogonal, denoting their lack of correlation. PCA can, for example, be used to reveal subpopulations of neurons characterized by similar activity patterns [83].

Delving into the PCA process, single-trial PSTHs of individual neurons are collected for each trial and are standardized to have a mean of zero and a standard deviation of one. This ensures that each variable contributes equally to the PCA, regardless of its original scale. Next, we compute the covariance matrix for the standardized single-trials PSTH data. Each row and column in this matrix represent different neurons (or channels), and entries within the matrix illustrate the degree of covariance—or shared variability—between the spike counts of two different neurons across all trials. A high covariance value would suggest that two neurons often fire in a correlated manner when subjected—for instance—to specific sensory stimuli. PCA is achieved by computing the eigenvectors and eigenvalues of the covariance matrix. Eigenvectors of this covariance matrix represent the directions along which the neural data varies the most, and the eigenvalues are proportional to the magnitude of that variance. Together, they form the "principal components" that serve as the data's new vector basis. Sorting and selecting the principal components are the next steps. These eigenvectors are sorted by the magnitude of their eigenvalues in descending order. The eigenvector with the highest eigenvalue becomes the first principal component, capturing the most significant variance within the data. This sorted set of eigenvectors forms a new orthogonal basis set for our data. Transitioning to the application phase, we employ these sorted eigenvectors to perform an orthogonal transformation of the original, standardized data. This transformation projects the high-dimensional neural data into a new orthogonal set of variables, the principal components (Fig. 8). Each principal component explains a unique level of variance, representing different patterns and structures in neural encoding. The final step in PCA is to evaluate how much variance each principal component explains. This helps us decide which principal components are most relevant for our analysis. This

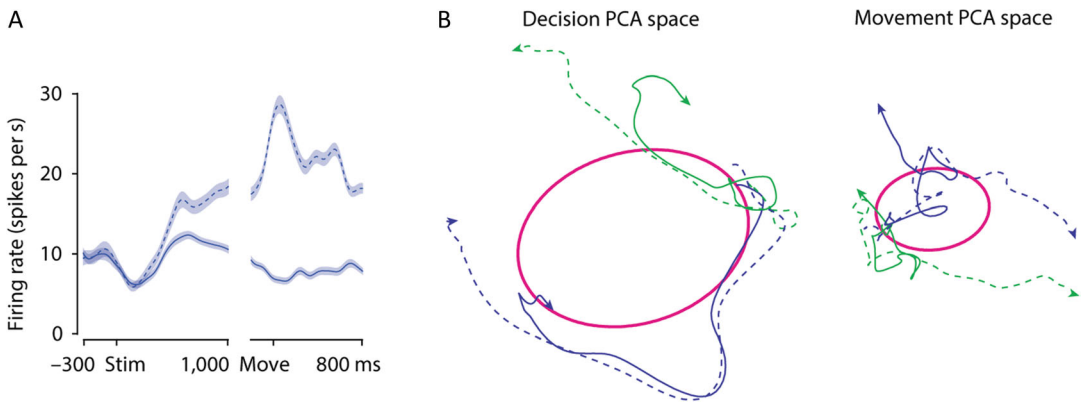

Fig. 8 Standard principal component analyses operations. (**a**): Time-varying firing rates of two example neurons during a stimulus-decision task. The first neuron exhibits a sustained preference for high-rate stimuli (represented by dashed lines) over low-rate ones (solid lines). The responses are aligned to stimulus onset on the left and movement onset on the right. This alignment is necessary due to the slight variation in time between the stimulus end and the animal's movement from trial to trial. Each subplot represents one neuron, with each line indicating one condition. The color codes the stimulus, while the solid/dashed line codes the binary decision. (**b**): Two-dimensional projections of decision-epoch data, consistent across all panels. These projections illustrate the variance in neuronal activity during the decision-making phase. Magenta ellipses indicate 1 standard deviation of the data when projected into the specified space. The spaces for these projections are chosen based on the principal components of the decision-epoch data. (Adapted with permission from Ref. [81])

selection process is based on the ration between the sum of the eigenvalues of the retained components and the sum of all eigenvalues. By focusing on the retained principal components, researchers can isolate the neuronal patterns associated with specific sensory inputs or behavioral outcomes, gaining a deeper understanding of the neural mechanisms at play in multisensory experiments.

PCA, while commonly employed in extracting dominant patterns from high-dimensional data, has certain inherent limitations. One of its primary constraints is its linearity. The method operates under the assumption that relationships among variables are linear. This can challenge capturing complex, nonlinear interactions often present in neural datasets [84]. Alternatives like Kernel PCA or t-distributed stochastic neighbor embedding (t-SNE) offer ways to capture nonlinear relationships and can be more suitable for such complex datasets.

Second, PCA's ranking of components by variance might only sometimes reflect the most biologically relevant features [85]. The method's sensitivity to data scale renders it vulnerable to outliers and noise, a notable concern in electrophysiology. Interpreting the biological significance of principal components, especially when they encapsulate a mix of neural sources, is often complex. Lastly, the orthogonality of principal components, a core assumption of

PCA, may sometimes align with the realities of biological processes. In multisensory research, essential cross-modal interactions could, for instance, be obscured if attention is limited to the dominant components.

3.3.3 Tensor Component Analysis (TCA)

Tensor component analysis (TCA) offers an alternative to traditional PCA for analyzing multidimensional neural data. Unlike PCA, which typically handles two-dimensional data matrices, TCA can manage higher-dimensional structures called tensors [82]. This enables TCA to capture patterns across dimensions like time, trials, and stimuli without requiring data preprocessing steps like concatenation. This feature makes TCA particularly useful in multisensory research. For example, in a study involving auditory and visual cues, TCA could analyze a four-way tensor encompassing time, neurons, trials, and additional factors such as pupil size, providing a more comprehensive view than PCA [14].

After collecting neural data across visual, auditory, and behavioral dimensions, we catalog each trial and time point to construct a foundational tensor (Fig. 9a). A tensor is a multidimensional generalization of a matrix, capable of representing complex data structures beyond just rows and columns. We then apply TCA to this tensor. The TCA method decomposes it into a core tensor and several factor matrices (Fig. 9b). The core tensor is a condensed representation capturing the original data's most significant relationships and variations. In contrast, the factor matrices offer granular insights into individual dimensions such as trials, sensory modalities, and time points. These matrices indicate the contribution of individual components identified by the TCA.

TCA allows us to identify specific latent dynamics associated, for example, with different sensory inputs. Unlike PCA, TCA retains the multidimensional nature of the data, making it easier to detect interactions among various sensory modalities, time points, and trials. For example, TCA can reveal how a group of neurons responds to visual stimuli and auditory cues and how these responses change over time. Visualization can map these complex interactions, providing a more holistic view of multisensory processing (Fig. 9c).

TCA provides a method for unsupervised analysis of latent dimensions underlying neural dynamics that is particularly useful in the context of multisensory integration. However, its application is not without limitations. A significant concern is the computational complexity associated with TCA. As the size and complexity of datasets increase, there is a corresponding rise in computational demands. This is inherent to TCA's capability to process multidimensional data; larger dimensions necessitate increased computational resources and time for effective analysis [86].

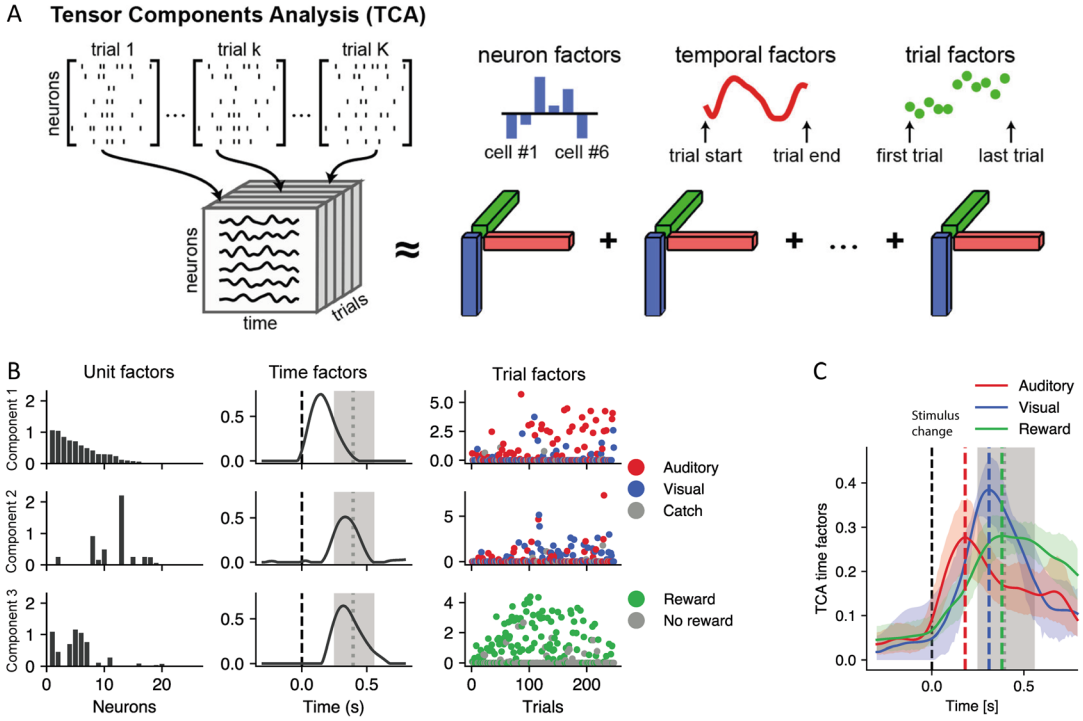

Fig. 9 TCA architecture and example of latent dimensions of parietal population activity correspond to task-relevant components. (**a**): TCA architecture developed by Williams et al. [82]. The dataset is structured into a third-order tensor with dimensions N × T × K. Utilizing TCA, the dataset is approximated by the summation of outer products derived from three orthogonal vectors. This methodology results in an ancillary set of low-dimensional factors, denoted as trial factors (represented in green), which systematically elucidate the temporal variations in neuronal activity across sequential trials. (**b**): Three selected components from a TCA analysis of an example recording session. Each row corresponds to a component (#1–3) that captures a subpopulation of neurons (unit factors, left column) that share a temporal response profile within the trial (temporal factors, central column), which is in turn expressed differently across trials (trial factors, right column). All factors are unitless. Neurons are ranked based on their contribution to the first plotted component, then the second. TCA is performed without information about trial types, colored a posteriori. Nonetheless, the identified latent components show selective expression in auditory (#1), visual (#2), and hit trials (#3). (**c**): Averaging the time factors of all selective components from different sessions showed a distinct temporal profile with a progression of first auditory, then visual, and last, decision components. (Adapted with permission from Refs. [14, 82])

Additionally, the interpretation of results derived from TCA presents challenges. The patterns identified by TCA, despite being multidimensional, may not always correspond directly to established or anticipated neural processes. This necessitates careful consideration when attempting to relate TCA-derived patterns to specific neural mechanisms [82].

3.3.4 Decoding

PCA and TCA are unsupervised methods, focusing on condensing and extracting dominant patterns from high-dimensional datasets without necessarily referring to specific outcomes. Decoding, on the other hand, adopts a supervised approach, aiming to predict specific experimental conditions or behavioral outcomes from patterns of neural activity. This distinction underscores a fundamental difference in goals: while unsupervised techniques aim for dimensionality reduction and feature extraction, supervised decoding is more outcome-oriented, making inferences about external variables based on neural activity.

Decoding methods come in various types, each with its own features. For example, support vector machines are used for classification and regression problems, excelling at finding the best dividing line or "hyperplane" that separates data into different categories [87]. Artificial neural networks are often applied to capturing complex, nonlinear dependencies across data sequences [88]. Finally, K-nearest neighbors is a supervised learning algorithm that classifies new data points by looking at their "k" closest neighbors in the training set and making a prediction based on the most common class or average value among these neighbors [87]. In this section, we will focus on another method: random forest decoding. One primary reason for this choice is its widespread use in neuroscience research since this approach offers high interpretability, can handle nonlinear interactions effectively, and is relatively robust to overfitting [89]. This technique is based on the use of multiple decision trees to make better predictions. A tree is a simple decision-making model that contributes its own "vote" on the outcome and, when combined with other trees, helps make the overall prediction more accurate and robust.

Delving into the process, initial steps such as data preprocessing and normalization are pivotal. Next, the random forest classifier is configured by constructing multiple decision trees. These trees work collectively, enhancing the predictive power and accuracy of the model. The complexity of this model is adjustable by tuning hyperparameters, such as the number of trees and their maximum depth. Finding an equilibrium is essential: while having more trees generally augments robustness, it simultaneously escalates computational requirements. Moreover, controlling the depth of the trees is vital to ensure our model neither underfits nor overfits the data.

In contrast with other decoding approaches, there is an ongoing debate in the field about the need to use cross-validation with random forest classifiers. Some researchers believe this is not needed, because the design of random forests naturally prevents overfitting, thanks to how it uses multiple decision trees. However, using cross-validation adds an extra check to ensure the model works well and is free from biases [90].

In the context of multisensory research, training the random forest involves mapping activity features, such as neural responses to

varying sensory stimuli, to a designated target variable. Once the model is trained, it employs a "majority vote" system among the individual trees for classification tasks or aggregates them for regression tasks. Essentially, each tree offers its independent prediction, with the model's final decision stemming from aggregating these numerous insights. This method is especially adept at managing high-dimensional data (Fig. 10a).

This type of decoder is often used to determine how well a set of experimental conditions (e.g., sensory stimulus) can be decoded over time. In this case, independent decoders may be trained at each time bin after—for instance—stimulus onset, to indicate how well neuronal activity in a given area tracks experimental variables (Fig. 10b).

The random forest algorithm, like all techniques, has some drawbacks. For instance, while seminal work on the algorithm highlighted its robustness and versatility, a subtle reading reveals concerns about interpretability [90]. Another aspect to consider is the sensitivity to hyperparameter settings. In machine learning, the careful choice of hyperparameters can differentiate between a successful and an ineffective model [92]. Finally, some studies provide insights into the performance of decoding algorithms, including random forests, across varying data landscapes [93].

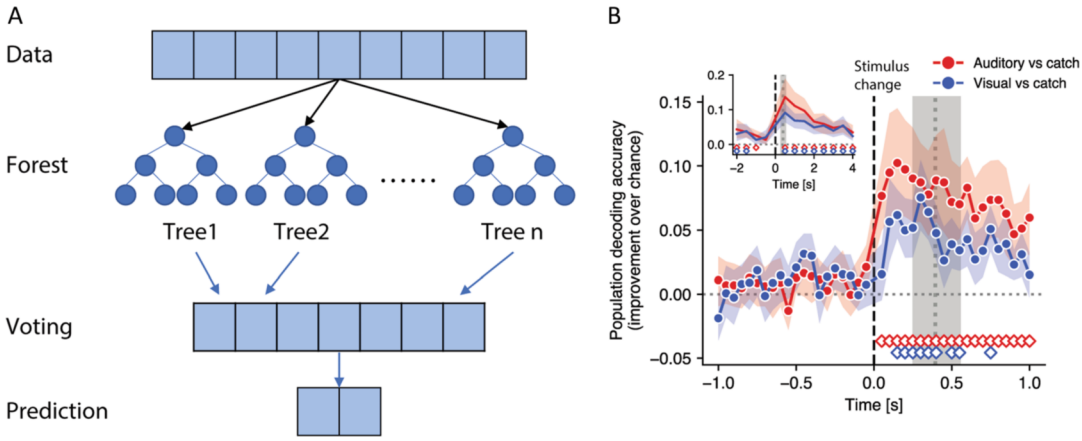

Fig. 10 Random forest architecture and example result: (**a**): Random forest architecture overview: the process begins with the dataset input, characterized by its features and observations. This leads to the construction of the decision trees. Each tree within this ensemble is crafted utilizing distinct subsets of data and features. As these trees process the data, they each arrive at their individual conclusions, culminating with the voting stage. After, this collective decision-making process ultimately conclude in the final prediction. (**b**): Accuracy of population decoding relative to chance plotted over time for random forest classifiers trained to discriminate visual versus catch trials (blue) and auditory versus catch trials (red). Trials were aligned to the time of stimulus change. Accuracy above chance was obtained by subtracting from the average decoding accuracy after repeatedly shuffling the trial labels. (Adapted with permission from Refs. [14, 91])

4 Conclusions

In this chapter, we presented an overview of the major experimental and analytical steps required to study multisensory processing in head-fixed behaving rodents, with a primary focus on silicon probe recordings. We especially paid attention to analysis methods aiming to understand how single neurons and neuronal populations (ensembles) process (multi)sensory information. Here we mainly focused on methods to analyze activity within brain regions. Several approaches are however emerging to analyze interactions between brain regions [94–96]. These will prove very valuable in the context of multisensory processing but are less established compared to the methodologies we surveyed here.

This focus stems from recent development of methods capable of simultaneously recording neuronal activity from hundreds of neurons located across cortical and subcortical regions, with action potential resolution [17, 44]. This remains currently unfeasible with other techniques, such as imaging [16], and is especially relevant for the multisensory field. In fact, multisensory processing has been observed in almost any brain region [6, 10]. However, recent studies have questioned whether all these supposedly multisensory signals actually have a different source (e.g., in terms of correlates of stereotypical motor responses to sensory stimuli) [22, 23]. Having the high-temporal resolution that only electrophysiology enables is essential to disambiguate true multisensory signals from other sources, using methods such as GLM regression models [23]. Nevertheless, all recording techniques can only yield intrinsically correlative insights about the link between brain activity and external factors. For this reason, integrating recording techniques with perturbational approaches such as optogenetics is essential to obtain mechanistic insights. This aspect goes beyond the scope of this chapter, but it must be pointed out that electrophysiology can be more easily integrated with optogenetics or other similar approaches, compared, for example, to two-photon imaging.

Compared to other, more established fields, multisensory processing is still in its infancy. While the basic mechanisms governing multisensory integration in subcortical structures such as the superior colliculus are well-known [7], much less is known about the cortex. In fact, not only the circuit-level mechanisms of cortical multisensory processing are still poorly understood [6, 9, 24], but there is even no consensus on which cortical areas actually process more than one modality [6, 10, 14, 22, 23, 97]. This chapter is meant as a guide for researchers willing to address this fascinating yet still relatively unexplored topic.

References

1. Spence C (2015) Multisensory flavor perception. Cell 161:24–35
2. Parr WV, Geoffrey White K, Heatherbell DA (2003) The nose knows: influence of colour on perception of wine aroma. J Wine Res 14:79–101
3. Gleiss S, Kayser C (2014) Acoustic noise improves visual perception and modulates occipital oscillatory states. J Cogn Neurosci 26:699–711
4. Iurilli G, Ghezzi D, Olcese U, Lassi G, Nazzaro C, Tonini R, Tucci V, Benfenati F, Medini P (2012) Sound-driven synaptic inhibition in primary visual cortex. Neuron 73:814–828
5. Chang S, Xu J, Zheng M, Keniston L, Zhou X, Zhang J, Yu L (2022) Integrating visual information into the auditory cortex promotes sound discrimination through choice-related multisensory integration. J Neurosci 42: 8556–8568
6. Meijer GT, Mertens PEC, Pennartz CMA, Olcese U, Lansink CS (2019) The circuit architecture of cortical multisensory processing: distinct functions jointly operating within a common anatomical network. Prog Neurobiol 174:1–15
7. Stein BE, Stanford TR (2008) Multisensory integration: current issues from the perspective of the single neuron. Nat Rev Neurosci 9:255–266
8. Angelaki DE, Gu Y, DeAngelis GC (2009) Multisensory integration: psychophysics, neurophysiology, and computation. Curr Opin Neurobiol 19:452–458
9. van Atteveldt N, Murray MM, Thut G, Schroeder CE (2014) Multisensory integration: flexible use of general operations. Neuron 81: 1240–1253
10. Ghazanfar AA, Schroeder CE (2006) Is neocortex essentially multisensory? Trends Cogn Sci 10:278–285
11. Wimmer RD, Schmitt LI, Davidson TJ, Nakajima M, Deisseroth K, Halassa MM (2015) Thalamic control of sensory selection in divided attention. Nature 526:705–709
12. Lee CCY, Diamond ME, Arabzadeh E (2016) Sensory prioritization in rats: behavioral performance and neuronal correlates. J Neurosci 36:3243–3253
13. Oude Lohuis MN, Pie JL, Marchesi P, Montijn JS, de Kock CPJ, Pennartz CMA, Olcese U (2022) Multisensory task demands temporally extend the causal requirement for visual cortex in perception. Nat Commun 13:2864
14. Oude Lohuis MN, Marchesi P, Pennartz CMA, Olcese U (2022) Functional (ir)relevance of posterior parietal cortex during audiovisual change detection. J Neurosci 42:5229
15. Carandini M, Churchland AK (2013) Probing perceptual decisions in rodents. Nat Neurosci 16:824–831
16. Machado TA, Kauvar IV, Deisseroth K (2022) Multiregion neuronal activity: the forest and the trees. Nat Rev Neurosci 23:683–704
17. Jun JJ, Steinmetz NA, Siegle JH et al (2017) Fully integrated silicon probes for high-density recording of neural activity. Nature 551:232–236
18. Olcese U, Iurilli G, Medini P (2013) Cellular and synaptic architecture of multisensory integration in the mouse neocortex. Neuron 79: 579–593
19. Deneux T, Harrell ER, Kempf A, Ceballo S, Filipchuk A, Bathellier B (2019) Context-dependent signaling of coincident auditory and visual events in primary visual cortex. elife 8. https://doi.org/10.7554/eLife.44006
20. Lippert MT, Takagaki K, Kayser C, Ohl FW (2013) Asymmetric multisensory interactions of visual and somatosensory responses in a region of the rat parietal cortex. PLoS One 8: e63631
21. Shams L, Kamitani Y, Shimojo S (2000) Illusions. What you see is what you hear. Nature 408:788
22. Bimbard C, Sit TPH, Lebedeva A, Reddy CB, Harris KD, Carandini M (2023) Behavioral origin of sound-evoked activity in mouse visual cortex. Nat Neurosci 26:251–258
23. Oude Lohuis MN, Marchesi P, Olcese U, Pennartz CM (2024) Triple dissociation of visual, auditory and motor processing in mouse primary visual cortex. Nat Neurosci 27(4): 758–771
24. Fetsch CR, DeAngelis GC, Angelaki DE (2013) Bridging the gap between theories of sensory cue integration and the physiology of multisensory neurons. Nat Rev Neurosci 14: 429–442
25. Pouget A, Deneve S, Duhamel J-R (2002) A computational perspective on the neural basis of multisensory spatial representations. Nat Rev Neurosci 3:741–747
26. Stringer C, Pachitariu M, Steinmetz N, Reddy CB, Carandini M, Harris KD (2019) Spontaneous behaviors drive multidimensional, brain-wide activity. Science 364:255
27. Steinmetz NA, Zatka-Haas P, Carandini M, Harris KD (2019) Distributed coding of

choice, action and engagement across the mouse brain. Nature 576:266–273
28. Zatka-Haas P, Steinmetz NA, Carandini M, Harris KD (2021) Sensory coding and the causal impact of mouse cortex in a visual decision. elife 10:e63163
29. Cunningham JP, Yu BM (2014) Dimensionality reduction for large-scale neural recordings. Nat Neurosci 17:1500–1509
30. Saxena S, Cunningham JP (2019) Towards the neural population doctrine. Curr Opin Neurobiol 55:103–111
31. Glasenapp A, Bankstahl JP, Bähre H, Glage S, Bankstahl M (2023) Subcutaneous and orally self-administered high-dose carprofen in male and female mice: pharmacokinetics, tolerability and impact on cage-side pain indicators. bioRxiv, 2023–06
32. Guo ZV, Hires SA, Li N et al (2014) Procedures for behavioral experiments in head-fixed mice. PLoS One 9. https://doi.org/10.1371/journal.pone.0088678
33. Tucci V, Hardy A, Nolan PM (2006) A comparison of physiological and behavioural parameters in C57BL/6J mice undergoing food or water restriction regimes. Behav Brain Res 173: 22–29
34. Urai AE, Aguillon-Rodriguez V, Laranjeira IC, Cazettes F, Laboratory TIB, Mainen ZF, Churchland AK (2021) Citric acid water as an alternative to water restriction for high-yield mouse behavior. eNeuro 8:ENEURO.0230. https://doi.org/10.1523/ENEURO.0230-20.2020
35. Pinto L, Koay SA, Engelhard B, Yoon AM, Deverett B, Thiberge SY, Witten IB, Tank DW, Brody CD (2018) An accumulation-of-evidence task using visual pulses for mice navigating in virtual reality. Front Behav Neurosci 12:36
36. Dombeck DA, Harvey CD, Tian L, Looger LL, Tank DW (2010) Functional imaging of hippocampal place cells at cellular resolution during virtual navigation. Nat Neurosci 13: 1433–1440
37. Keller GB, Bonhoeffer T, Hübener M (2012) Sensorimotor mismatch signals in primary visual cortex of the behaving mouse. Neuron 74:809–815
38. Green DM, Swets JA (1966) Signal detection theory and psychophysics. Wiley, New York
39. Sridharan D, Steinmetz NA, Moore T, Knudsen EI (2014) Distinguishing bias from sensitivity effects in multialternative detection tasks. J Vis 14:1–32
40. Nicolai EN, Michelson NJ, Settell ML, Hara SA, Trevathan JK, Asp AJ, Stocking KC, Lujan JL, Kozai TDY, Ludwig KA (2018) Design choices for next-generation Neurotechnology can impact motion artifact in electrophysiological and fast-scan cyclic voltammetry measurements. Micromachines (Basel) 9:494
41. Zhuang J, Ng L, Williams D, Valley M, Li Y, Garrett M, Waters J (2017) An extended retinotopic map of mouse cortex. elife 6:e18372
42. Juavinett AL, Nauhaus I, Garrett ME, Zhuang J, Callaway EM (2017) Automated identification of mouse visual areas with intrinsic signal imaging. Nat Protoc 12:32–43
43. Steinmetz NA, Koch C, Harris KD, Carandini M (2018) Challenges and opportunities for large-scale electrophysiology with Neuropixels probes. Curr Opin Neurobiol 50:92–100
44. Steinmetz NA, Aydin C, Lebedeva A et al (2021) Neuropixels 2.0: a miniaturized high-density probe for stable, long-term brain recordings. Science 372:eabf4588
45. Angotzi GN, Boi F, Lecomte A, Miele E, Malerba M, Zucca S, Casile A, Berdondini L (2019) SiNAPS: an implantable active pixel sensor CMOS-probe for simultaneous large-scale neural recordings. Biosens Bioelectron 126:355–364
46. Neural Probes | Cambridge NeuroTech (2023). https://www.cambridgeneurotech.com/neural-probes. Accessed 25 Jul 2023
47. Electrode Arrays – NeuroNexus (2023). https://neuronexus.com/products/electrode-arrays/. Accessed 25 Jul 2023
48. Hong G, Lieber CM (2019) Novel electrode technologies for neural recordings. Nat Rev Neurosci 20:330–345
49. Senzai Y, Fernandez-Ruiz A, Buzsáki G (2019) Layer-specific physiological features and Interlaminar interactions in the primary visual cortex of the mouse. Neuron 101:500–513.e5
50. Juavinett AL, Bekheet G, Churchland AK (2019) Chronically implanted Neuropixels probes enable high-yield recordings in freely moving mice. elife 8:e47188
51. Luo TZ, Bondy AG, Gupta D, Elliott VA, Kopec CD, Brody CD (2020) An approach for long-term, multi-probe Neuropixels recordings in unrestrained rats. elife 9:e59716
52. van Daal RJJ, Aydin Ç, Michon F, Aarts AAA, Kraft M, Kloosterman F, Haesler S (2021) Implantation of Neuropixels probes for chronic recording of neuronal activity in freely behaving mice and rats. Nat Protoc 16:3322–3347
53. Okun M, Lak A, Carandini M, Harris KD (2016) Long term recordings with immobile silicon probes in the mouse cortex. PLoS One 11:e0151180

54. Connors BW, Gutnick MJ (1990) Intrinsic firing patterns of diverse neocortical neurons. Trends Neurosci 13:99–104
55. McCormick DA, Connors BW, Lighthall JW, Prince DA (1985) Comparative electrophysiology of pyramidal and sparsely spiny stellate neurons of the neocortex. J Neurophysiol 54: 782–806
56. Nowak LG, Azouz R, Sanchez-Vives MV, Gray CM, McCormick DA (2003) Electrophysiological classes of cat primary visual cortical neurons in vivo as revealed by quantitative analyses. J Neurophysiol 89:1541–1566
57. Kvitsiani D, Ranade S, Hangya B, Taniguchi H, Huang JZ, Kepecs A (2013) Distinct behavioural and network correlates of two interneuron types in prefrontal cortex. Nature 498: 363–366
58. Fiorilli J, Marchesi P, Ruikes T, Huis in 't Veld G, Buckton R, Duque Quintero M, Reiten I, Bjaalie J, Pennartz CM (2023) Neural correlates of object identity and reward outcome in the corticohippocampal hierarchy: double dissociation between perirhinal and secondary visual cortex. bioRxiv (2023): 2023-05
59. Pachitariu M, Steinmetz N, Kadir S, Carandini M, Kenneth DH (2016) Kilosort: realtime spike-sorting for extracellular electrophysiology with hundreds of channels. BioRxiv061481
60. Pachitariu M, Sridhar S, Stringer C (2023) Solving the spike sorting problem with Kilosort. BioRxiv2023.01.07.523036
61. Rossant C, Kadir SN, Goodman DFM et al (2016) Spike sorting for large, dense electrode arrays. Nat Neurosci 19:634–641
62. Yger P, Spampinato GL, Esposito E et al (2018) A spike sorting toolbox for up to thousands of electrodes validated with ground truth recordings in vitro and in vivo. elife 7: e34518
63. IronClust (RRID:SCR_018903), https://github.com/flatironinstitute/ironclust
64. Buccino AP, Hurwitz CL, Garcia S, Magland J, Siegle JH, Hurwitz R, Hennig MH (2020) SpikeInterface, a unified framework for spike sorting. elife 9:e61834
65. Lewicki MS (1998) A review of methods for spike sorting: the detection and classification of neural action potentials. Network 9:R53–R78
66. Tolias AS, Ecker AS, Siapas AG, Hoenselaar A, Keliris GA, Logothetis NK (2007) Recording chronically from the same neurons in awake, behaving primates. J Neurophysiol 98:3780–3790
67. Introduction – phy (2023). https://phy.readthedocs.io/en/latest/. Accessed 2 Sep 2023
68. Wen H, Shi J, Zhang Y, Lu K-H, Cao J, Liu Z (2018) Neural encoding and decoding with deep learning for dynamic natural vision. Cereb Cortex 28:4136–4160
69. Xu L, Xu M, Jung T-P, Ming D (2021) Review of brain encoding and decoding mechanisms for EEG-based brain–computer interface. Cogn Neurodyn 15:569–584
70. Barack DL, Krakauer JW (2021) Two views on the cognitive brain. Nat Rev Neurosci 22:359–371
71. Oude Lohuis MN, Canton AC, Pennartz CMA, Olcese U (2021) Higher order visual areas enhance stimulus responsiveness in mouse primary visual cortex. Cereb Cortex 32(15):3269–3288
72. Reches A, Gutfreund Y (2009) Auditory and multisensory responses in the tectofugal pathway of the barn owl. J Neurosci 29:9602–9613
73. Stevenson RA, Ghose D, Fister JK, Sarko DK, Altieri NA, Nidiffer AR, Kurela LR, Siemann JK, James TW, Wallace MT (2014) Identifying and quantifying multisensory integration: a tutorial review. Brain Topogr 27:707–730
74. Knöpfel T, Sweeney Y, Radulescu CI, Zabouri N, Doostdar N, Clopath C, Barnes SJ (2019) Audio-visual experience strengthens multisensory assemblies in adult mouse visual cortex. Nat Commun 10(1):5684
75. Movahedi F, Padman R, Antaki J (2021) Limitations of receiver operating characteristic curve on imbalanced data: assist device mortality risk scores. J Thorac Cardiovasc Surg 165: 1433. https://doi.org/10.1016/j.jtcvs.2021.07.041
76. Poline J-B, Brett M (2012) The general linear model and fMRI: does love last forever? NeuroImage 62:871–880
77. Meijer GT, Marchesi P, Mejias JF, Montijn JS, Lansink CS, Pennartz CMA (2020) Neural correlates of multisensory detection behavior: comparison of primary and higher-order visual cortex. Cell Rep 31:107636
78. Vaccari FE, Diomedi S, Filippini M, Hadjidimitrakis K, Fattori P (2022) New insights on single-neuron selectivity in the era of population-level approaches. Front Integr Neurosci 16:929052
79. Kobak D, Brendel W, Constantinidis C, Feierstein CE, Kepecs A, Mainen ZF, Qi X-L, Romo R, Uchida N, Machens CK (2016) Demixed principal component analysis of neural population data. elife 5:e10989
80. Gallego JA, Makin TR, McDougle SD (2022) Going beyond primary motor cortex to improve brain–computer interfaces. Trends Neurosci 45:176–183

81. Raposo D, Kaufman MT, Churchland AK (2014) A category-free neural population supports evolving demands during decision-making. Nat Neurosci 17:1784–1792
82. Williams AH, Kim TH, Wang F, Vyas S, Ryu SI, Shenoy KV, Schnitzer M, Kolda TG, Ganguli S (2018) Unsupervised discovery of demixed, low-dimensional neural dynamics across multiple timescales through tensor component analysis. Neuron 98:1099–1115.e8
83. Ringnér M (2008) What is principal component analysis? Nat Biotechnol 26:303–304
84. Vaccari FE, Diomedi S, Filippini M, Galletti C, Fattori P (2021) A Poisson generalized linear model application to disentangle the effects of various parameters on neurophysiological discharges. STAR Protocols 2:100413
85. Björklund M (2019) Be careful with your principal components. Evolution 73:2151–2158
86. Kolda TG, Bader BW (2009) Tensor decompositions and applications. SIAM Rev 51(3): 455–500
87. Fawzy H, Rady EHA, Abdel Fattah AM (2020) Comparison between support vector machines and k-nearest neighbor for time series forecasting. J Math Comput Sci 10(6):2342–2359
88. Zhang H, Li L, Li X (2022) Exponential synchronization of coupled neural networks under stochastic deception attacks. Neural Netw 145: 189–198
89. Schooten SV, Harel R, Ercan S, Groot ED (2014) Applying feature selection methods on fMRI data
90. Breiman L (2001) Random forests. Mach Learn 45:5–32
91. Kong Y, Yu T (2018) A deep neural network model using random forest to extract feature representation for gene expression data classification. Sci Rep 8:16477
92. Probst P, Wright MN, Boulesteix AL (2019) Hyperparameters and tuning strategies for random forest. Wires Data Min Knowl J 9(3): e1301
93. Hastie T, Tibshirani R, Friedman JH, Friedman JH (2009) The elements of statistical learning: data mining, inference, and prediction, vol 2. springer, New York, pp 1–758
94. Gokcen E, Jasper AI, Semedo JD, Zandvakili A, Kohn A, Machens CK, Yu BM (2022) Disentangling the flow of signals between populations of neurons. Nat Comput Sci 2:512–525
95. Semedo JD, Jasper AI, Zandvakili A, Krishna A, Aschner A, Machens CK, Kohn A, Yu BM (2022) Feedforward and feedback interactions between visual cortical areas use different population activity patterns. Nat Commun 13:1099
96. Semedo JD, Gokcen E, Machens CK, Kohn A, Yu BM (2020) Statistical methods for dissecting interactions between brain areas. Curr Opin Neurobiol 65:59–69
97. Pennartz CMA, Oude Lohuis MN, Olcese U (2023) How "visual" is the visual cortex? The interactions between the visual cortex and other sensory, motivational and motor systems as enabling factors for visual perception. Philos Trans R Soc Lond Ser B Biol Sci 378: 20220336