

Cortex-wide imaging in naturally behaving mice with 1-gram wireless mesoscopes

Corresponding Author: Professor Qionghai Dai

Parts of this Peer Review File have been redacted as indicated to maintain the confidentiality of unpublished data.

Parts of this Peer Review File have been redacted as indicated to remove third-party material.

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Referee #1

(Remarks to the Author)

The manuscript by Zhang and coworkers describes 'Airscope', a wireless mesoscope epifluorescence system featuring a 6 mm diameter field-of-view with lateral 4 – 5 micrometer optical resolution and 10 Hz frame rate, weighting 1 g. The 1.7 g supplied battery enables 15 minutes autonomy and can be rapidly swapped to continue recording. An alternative wired power line enables long uninterrupted recordings. The authors describe a fully automated analysis pipeline featuring SAM2MICE, a behaving-mice segmentation algorithm for automated videography analysis based on foundational model SAM2. Together these techniques are used in a wide range of behavioral paradigms including social interaction, swimming and large environment exploration. These impressive technical advances enable a robust and versatile device, which linked to an open-sourced release promises to have an impact in neuroscience of freely-moving animals if a few of the details of the microscopes capabilities and caveats are demonstrated fully.

While I applaud the authors on their engineering feat and potential application of the microscope, at present many of the major claims of the paper are somewhat unjustified as they are presented and need addressing. Overall, the manuscript is very unfocussed and moves from the new technique to behavior to brain activity correlates and behavior. While I understand that the aim was to show the many different potential applications that the new approach could be applied to, the issue is that it is hard to understand if any of the biological claims are true and whether they would hold up to deeper scrutiny. This leads to the biological results being very descriptive and relying on 'black-box' analysis techniques to make any sense of the data but leaves the reader unable to really assess the quality of the raw data being gathered by the microscope and the subsequent step by step analysis steps.

Major:

1) If optics only are considered, previous work by Sherrer et al. already published an optical system based on injection-molded aspheric lenses. This previous system is well acknowledged by authors and a quantitative comparison is available, presenting an improvement by the new design (factor 2 of space-bandwidth metric), which overall is an incremental advance and should be presented as a solid step but not a major innovation and certainly was not the first to apply this approach to lightweight 1P microscopes.

2) The space-bandwidth metric is well suited to evaluate optics, but does not include temporal resolution. In the same previous work, the mesoscope operated at 30 Hz frame rate, whereas Airscope performs at 10 Hz. A more revealing of overall performance of the whole system (not just optics) would be a time-space-bandwidth metric, which would be relatively close between the kiloscope from Sherrer et al and Airscope.

3) The single-cell resolution claimed in the manuscript is unjustified. Cellular resolution is in 3D not just 2D, and while the lateral resolution is enough for lateral resolution this is not the case in the axial plane. Given the tight packing of neuronal somas in the cortex in both lateral and axial planes, how can the activity from neurons stacked in the axial plane be separated? This is especially important when considering the biological claims later in the manuscript where single-cell-resolution are required.

4) The claim of the number of neurons imaged seems somewhat difficult to reliably measure with 1P excitation. The claims seem to be based on how many soma-like shapes can be detected but not based on the reality of how many neurons are present in the tissue, or even how many somas are labelled vs not labelled, i.e. the density of labelling. Given that the techniques used to label the tissue what fraction of the total number of neurons in the area did you image, only the brightly labelled neurons? How was it determined that cortical layer 2/3 are being imaged? Please provide some ground-truth to the imaging claims of number of neurons imaged in each of the experiments.

5) It is not clear throughout the manuscript what the benefits of the system are compared to tethered system, it would be important to show an experiment from the same brain-tissue where activity maps could be showed using the untethered microscope that was different to activity maps when using the tethered scope. As it stands, the main argument is that tether free is better because of the application rather than a finding. This is ok, but with the lack of single-cell resolution and the lack of controls, especially controls for linking the correlated activity to behavior, there are observations of neuronal activity but no causative controls. This makes the biological claims somewhat difficult to understand.

6) There are many examples of overclaiming for cortical function in which the data does not stand up. I will give a few examples, but these are not by any means the only examples, please go through the manuscript and remove or rewrite these to fit the data accurately.

From the abstract:

39 Using Airscope, we provide an in-
40 itial appraisal of untethered brain-behavior dynamics.

Im not sure what an initial appraisal is in this context, especially in the abstract.

42 comes across animals. We identify a previously uncharacterized reward discovery state
43 during maze exploration.

I cannot find the data and control experiments that literally show this finding beyond correlation.

43 We map a new causal pathway from spinal cord to the cortex
44 causal flow during fear conditioning enabled by simultaneous brain-spinal recordings.

I cannot find data that shows a causal pathway. This implies anatomy and causal experiments.

437 Airscope unveiled cortical basis of animals' natural exploration in a labyrinth

How can this be true? If only layer 2/3 were being imaged, then what is the rest of the cortex doing during this task? Where are the control experiments? Where is the raw data that shows this effect? How does the model work? Saying that the cortical basis of an animal's maze exploration has been unveiled is literally not true.

7) Open-source access to these technologies would maximize the impact of this advance on the field. But to be open source, the authors need to fully release electronic and optical designs. Crucially, Zemax designs must be made available for future improvements of the system and to give scientists that might use the system in the future the full understanding of the data generated with this device.

8) The discussion is weak and missing many many key references, especially in light of the head-mounted technology and biological claims in the manuscript. It would be important to discuss the microscope in terms of the other head-mounted microscopes that employ 3 and 2-photon excitation and provide optical sectioning. All microscopes have both pros and cons over the current microscope, these should be discussed succinctly and in a balanced approach.

In another example, the current method attempts to scene reconstruction by adding another camera to the head, this has been done before and these papers should be cited. Scene reconstruction of freely moving mice has been achieved by two methods, one by a head mounted camera (Parker et al 2022 Joint coding of visual input and eye/head position in V1 of freely moving mice) and by eye-tracking and modeling (Holmgren 2020 Visual pursuit behavior in mice maintains the pursued prey on the retinal region with least optic flow).

Minor aspects:

Authors report uniform 4 um resolution, but it does seem to degrade on the MTFs towards the edges? The authors extensively quantify the resolution and report correctly the values in the result paragraph (l.219-232), but later only mention central resolution, while it clearly degrades towards periphery (ex. l.244, table1 only 4 um value is reported). Upper bound or

range should be used instead.

L623-624 “optical design that leverages aspherical lens systems to improve per gram space -bandwidth product over 50 times”

This statement is confusing. By comparing to optics in the closest competitor, that also used aspheric lens system the improvement factor estimated by the authors is 2 and normalized by weight 2.8 (although this is weight of full system, not just optics). Factor 50 compares with closest wireless competitor, but that is irrelevant to optics.

The true outstanding advance here is the wireless transmission, that maintains sufficiently high data rate and is implemented in a lightweight and compact footprint. The authors show extensive application possibilities enabled by the absence of wires in a truly impressive set of behavioral paradigms. But it somehow lacks one clear example, where this key innovation giving access to completely unrestrained behavior enables new types of scientific questions. I do not mean here a niche experiment paradigm, like swimming or sheltering, but rather experiments enabling novel approaches to general neuroscience questions.

There are many weird sentence constructions that do not make sense such as:

444 Here, to our knowledge, we used the Airscope system to measure cortical activity during

445 large-scale labyrinth navigation (Fig. 5a and Supplementary Fig. 19a-c).

If possible, to show a system with a faster imaging rate would significantly enhance the capability of the system, even at the cost of some additional weight or in a wired mode in case the wireless transmission is the limiting step for faster imaging. Providing Zemax files would give this work the possibility to become a reference and benchmark for all future designs and would certainly maximize impact.

(Remarks on code availability)

Referee #2

(Remarks to the Author)

In this manuscript, a new miniature wireless one-photon microscope with large field of view is developed and used in mice as they engage in various behaviors where imaging with a tethered miniature microscope or a tabletop microscope would be difficult or impossible. In addition to the basic introduction of the new technology, the utility and performance of this new device is showcased during exploration in a naturalistic habitat and various other setups. Together with the hardware, multiple software tools are developed, including for neuron segmentation, behavior tracking, decoding activity, and other purposes. The hardware development on the side of optics, onboard processing, wireless transmission, and integration across systems is impressive.

Overall, this is a very impressive suite of work, that likely will be highly impactful for systems neuroscience, including for recording neural activity while mice explore in outside spaces.

In addition to this enthusiastic overall assessment, below are a number of comments and concerns:

In my opinion, the text and figures somewhat underemphasize the strength of the technology development and could provide readers new to the field with more intuition about how this progress was possible. For example, on the side of optics, it is described only briefly why the Airscope optical design allows for large field of view imaging using a short optical path length, and explains little about the relationship to the Chevalier landscape lens, how it was manufactured, and so on. These sections do not need to be longer, but they could be made more accessible to non-expert readers.

The demonstrations of the utility of the technology in the main text are extensive, but the characterization of the microscope itself is not as extensive and is covered in more detail in the Supplementary Material. It would be useful to move some of that characterization to the main text. Especially the axial resolution, measured in Supplementary Figure 10b, is useful to include, because it shows that signals from neurons at different depths can overlap and have to be recovered using algorithms like DeepWonder, used in the manuscript. The larger FWHM axially compared to laterally is expected as this is a 1-photon system, and to my knowledge it is often not shown in 1-photon miniscope publications, so it is good that the authors characterized it here.

A recurrent (but solvable) issue with the manuscript is the slight but pervasive over-interpretation of the biological results, as occurs for almost all experiments. For instance, in the decoding of social hierarchy, the manuscript claims that there is a cortex-wide representation of hierarchy, and posits that a previous result suggesting this was encoded in prefrontal cortex is incomplete. While it is plausible that indeed these representations are cortex-wide, the evidence presented does not show this. That is because the neural activity may contain correlates of the different types of motor actions performed by the animals, which are different for the dominant and subordinate mice. For example, the dominant mouse probably pushes more and the subordinate mouse walks backward more, allowing for purely motor representations to be informative about social rank, even if they do not explicitly encode rank. To claim that the cortical representations encode place in the social hierarchy would require very careful disambiguation from differences in motor actions and sensory experiences. Therefore, this claim should be toned down.

A similar comment applies to the experiment with the social feeding setup. Here, the cortical activity patterns that allow social status to be decoded may be neurons representing eating or taste, for example. Similarly, the claims are overstated as they are written now, and should be toned down.

For the landscape exploration decoding, where different zones in the landscape can be decoded from cortical activity, the same comment applies. Whether these are encoding 'landscape-specific cortical representations' in a more abstract way, or something lower level like average color, different spatial (visual) frequencies, or possibly location, cannot be deduced without careful control experiments. Again, the claims should be changed to what can actually be concluded. There are other cases throughout the manuscript where the biological claims should be revisited and possibly reworded in a more concrete way.

Importantly, these above comments about over-interpretation of the data are not negative, because nobody would expect a single methods-focused study to discover all these cortical representations. It is already impressive that these task and behavioral variables can be decoded from the data using a single field of view, and an adequate demonstration of the utility of the technology.

Layer specificity: the microscope is designed to record from a fixed depth and has, to my understanding, no option to change the imaging plane after mounting. This should be clearly stated. It does raise the question of whether there is a mechanism to adjust the vertical position before fixing the microscope, or whether the setup can be extended with an integrated simple vertical displacement such as a miniature stage or screws to adjust the height afterwards. It would be useful to at least discuss this.

That the manuscript offers not only a novel microscope but also a set of developed algorithms that can be used to get the most benefit from the microscope can be described more clearly. For example, the authors could create an integrated package that brings the behavioral and image analysis together in one place for increased adoption of the entire suite of methods.

Some comments on the figures for clarity. Figure 1a is confusing because it is unclear what the arrows mean. Readers might assume, for example, that the first panel represents the implanting of the Airscope, the middle panel a wired mode, and the third panel the wireless mode. It should be clarified that the first two panels are not what the manuscript is about. Regarding panel 1m, the background appears to be replaced by a color gradient, which at first made me think this was an illustration instead of a real photograph. Can the legend clarify that the background was replaced, or can the full photograph be shown?

In Figure 2b and 2c, the first two images showcasing the quality of the Airscope show blood vessels and microglia. This seems like an odd choice, given that the primary use of the Airscope is to image neural activity. A raw image of neurons is shown in 2f, but it is smaller and later. It seems that for a paper focused on neurons, neuron tissue should be shown first and in full detail.

The GitHub link, https://github.com/Neurallabware/Airscope_toolbox, appears not to work, so this code could not be reviewed. Perhaps the repository has not been set to public yet?

The optics, PCB board designs, firmware, and so on are available as open source as listed under the Data availability section. This is great but it is not immediately clear because only "CAD models" is mentioned. It would be beneficial to showcase more clearly what is in the google drive. Can you list the items so readers know that the full design is there?

Minor comments

The abstract contains some language that is hard to understand. "track length" (line 35) should be clarified to mean optical path length from the sensor to the edge. "tabletop-grade throughput" should be clarified to mean the number of neurons recorded using a tabletop one-photon microscope.

Line 52, "billions of neurons": this seems to refer to larger brains, such as the human brain, but is mentioned without specifying this. That is confusing since the manuscript is about mouse brains, which contain fewer neurons.

Line 69, "Even" should probably be "Even though".

The cost of 2 dollars mentioned on line 128, is this only for the lens array or for the entire device? It must be for the lenses only, since the sensor and chips likely cost more - please clarify this.

Line 269, "monthly" should probably say "month-long".

Line 445, "to our knowledge" appears to have been inserted in error.

Multiple axes and color bars are unlabeled and only defined in the legend, for example Fig. 6m and 6r. There is space to add labels and it would make the figures easier to interpret. The same comment applies to a number of other figure panels.

(Remarks on code availability)

The GitHub address didn't work - I am guessing the authors forgot to make the repository public.

Version 1:

Reviewer comments:

Referee #1

(Remarks to the Author)

The authors have made substantial effort to comprehensively answer all of the reviewer comments. They agreed with the reviewers about optical characterization metrics, including the spatio-temporal bandwidth, provided all requested technical files and software for open access, adjusted the biological claims. The result is a substantially clearer paper with sharper quantifications and validation of the very impressive technical advance described here. I have no more concerns and thank the authors for the upfront discussion. Well done.

(Remarks on code availability)

Referee #2

(Remarks to the Author)

The authors performed an extremely thorough and high-quality revision. My comments have been comprehensively addressed, and I have no further feedback apart from expressing my enthusiasm for this work.

(Remarks on code availability)

I have inspected but not run the code, as this would be the appropriate prior infrastructure I do not currently possess.

Referee #3

(Remarks to the Author)

The manuscript presents an impressive technological advance. Airscope combines large field of view, low weight and wireless operation in a way that should enable experiments difficult with existing head-mounted approaches. The additional validation is extensive, and I do not see a need for further experimental work except potentially on one point concerning the spinal cord recordings.

Major comment:

The manuscript convincingly establishes single neuron resolution in cortex through comparison with two-photon ground truth. However, I am less convinced that equivalent functional single-neuron resolution has been demonstrated in the spinal cord.

In the spinal experiments, the authors explicitly state that strong scattering necessitates the use of CellSort to extract four functional components rather than segmenting individual neurons. Figure 6 likewise describes these signals as four spinal cord components. The reported approximately 4 μm optical resolution is impressive, but optical resolution alone does not demonstrate that functional calcium signals can be reliably assigned to individual spinal neurons during freely moving recordings.

As currently presented, I would therefore interpret these data as spatially resolved spinal population or ensemble activity rather than validated single-neuron imaging. This distinction is important because spinal imaging is presented as a major capability of the platform. The authors need to be more explicit about single neurons resolution capabilities in the spinal cord or lack thereof.

If the authors intend to claim population level spinal imaging, I do not think further experiments are necessary, but the terminology should be revised accordingly. If, however, single neuron functional spinal imaging is intended as a central claim, I think direct experimental validation would be required, for example through sparse labelling or an independent cellular-resolution ground truth.

Minor comments:

1. The new transection experiment substantially strengthens the manuscript. Cortical responses to hindpaw stimulation are reduced almost to baseline after complete spinal transection.

This convincingly demonstrates that intact ascending spinal transmission is necessary for the cortical response to hindpaw stimulation. It does not, however, establish that the specific spinal components recorded by Airscope causally drive cortical activity, nor does it directly establish a spinal mechanism for the learned auditory cue response during fear conditioning. I would therefore keep the causal interpretation tightly restricted to the shock-evoked ascending pathway.

2. The revised term "Granger-causal influence" is appropriate, provided it remains clearly distinguished from anatomical or interventional causation. I would ask the authors to clarify the analysis window, whether cue and shock epochs were treated separately, how trials were concatenated, whether movement or stimulus timing were included as covariates, and how stationarity was assessed. No additional experiment is required.

3. The effect is large and convincing, but the reported significance for single-neuron responses uses >4k neurons from only three animals. Since neurons within an animal are not independent biological replicates, statistical inference should preferably use the animal as the experimental unit or an appropriate hierarchical analysis.

4. The cortical recordings primarily sample layer 2/3 across dorsal cortex rather than the whole brain. I would therefore favour "dorsal cortex-wide" or "hemisphere-wide cortical" rather than "whole-brain." Similarly, unless cellular identity is

directly demonstrated in the spinal preparation, terms such as "spinal population activity", "functional components" or "ensembles" would be more precise than "spinal neurons."

Recommendation

I would be comfortable with the manuscript moving forward once the language is adjusted as suggested to more faithfully reflect the experimental evidence. The only point that may require further experimental validation is the claim of single-neuron functional imaging in the spinal cord. If that claim is restricted to spatially resolved population imaging, and this is made clear in the manuscript, I do not think additional experiments are necessary.

(Remarks on code availability)

Referee #4

(Remarks to the Author)

The manuscript by Y. Zhang et. al. entitled "Deciphering cortex-wide neural dynamics of naturally behaving mice by 1-gram wireless mesoscope" reports a miniaturized microscope (termed as Aircscope) that can be head-mounted on the mouse head to record its brain activity while the animal freely-behaves. The Aircscope can be controlled wirelessly and transmit the imaging data wirelessly, which allows untethered imaging of mouse cortex and spinal cord. The Aircscope weighs 1.0 gram, has 4-6 μm lateral resolution across a 6 mm field of view, and can perform recording at 12.5 Hz. Using the Aircscope, the authors demonstrated neuronal activity recording under a variety of animal behaviors, including multi-animal free exploration over a 5 m² arena, competitive interaction between two animals, water reward searching in labyrinth, and response to pain by foot shocks and fear conditioning.

The results reported in the manuscript are very impressive. From the technical aspects, Aircscope outperforms existing one-photon miniaturized microscopes, in particular those with wireless recording capability, in terms of spatial bandwidth product and spatiotemporal bandwidth product. In particular, the Chevalier architecture with few lenses design and the wireless transmission module with on-board data compression is quite innovative. The multi-mice tracking algorithm (SAM2Mice) also represents a very powerful tool. From the systems neuroscience perspective, Aircscope enables a wide range of studies that were challenging to perform in the past. Large-scale calcium imaging can now be performed across multiple animals over a large arena without concerning the tethered wires tangling together. The biological findings, mostly about how neuronal activity encodes a specific animal behavior, are quite interesting. The compact footprint, light weight and wireless features of the Aircscope should reduce the impact of the head-mounted device on the animals, and thus make the animal behavior more naturalistic. I believe Aircscope represents a significant advancement in the field, and it can be widely adopted by the systems neuroscientists. The results presented in the paper are comprehensive and the statistical tests are appropriate. It has the merit to get published in Nature.

While the revised manuscript has addressed the previous reviewers' comments and is of very high quality, it could be further strengthened by the consideration of the following comments.

1. One bottleneck of the previous wireless miniscopes is the limited transmission bandwidth. Aircscope compresses the data prior to transmission. While there is a supplementary figure to describe this, the description is very sparse in the main text. In particular, the detail of the implementation is missing in Methods. More details on how to record the wireless data from multiple animals simultaneously should also be added in Methods.
2. There is a lack of discussion about the thermal management of the device. In particular, what is the total heat generated of the device on the animal?
3. Regarding the LED illumination, is there a collimation lens after the LED? This needs to be clarified.
4. The authors developed SAM2Mice for multi-animal tracking. The authors also previously developed an "unsupervised deep transfer learning method for multi-animal tracking (UDMT)" for a similar task. It will be good to comment on how SAM2Mice is compared with UDMT.
5. For the competitive interaction experiments shown in Fig. 4, how many (or percentage of) neurons are selective for push, resist and retreat? When these neurons were removed from the decoder input, how many neurons were left? One interesting question is to evaluate how many neurons are needed to encode/predict the trial outcome. Such a question can only be answered through devices like Aircscope. I would like to mention that such a result is interesting to see, but it is ok if the authors decide not to include in this manuscript, particularly when the results are already very dense in the manuscript.

Below are some minor comments:

6. Lensless approach is one competing technique to achieve large scale imaging with small footprint. The authors mentioned this category in the introduction. One most recent advance in this field is the following paper.
<https://www.science.org/doi/full/10.1126/sciadv.adr6687>

7. In Supplementary Table 1, the authors may want to add a column to indicate whether this is 1P or 2P/3P miniscope, or separate the table into two sessions for 1P, and 2P/3P. It will be more comprehensive to include the recently developed beam-multiplexed 2P/3P miniscope, which has an increased spatiotemporal bandwidth product, such as

[https://www.cell.com/cell-reports-methods/fulltext/S2667-2375\(25\)00257-7](https://www.cell.com/cell-reports-methods/fulltext/S2667-2375(25)00257-7)
<https://www.nature.com/articles/s41592-026-03125-7>

8. Fig. 1 m-o, please include the recording time in the caption.
9. Supplementary Fig. 4d, quantitative value on the vertical axis is missing.
10. Supplementary Fig. 5, does this assembly include the option of SD card placement?

(Remarks on code availability)

I went through the above github repo as well as the two other websites provided in the manuscript:

<https://airscope.org/devkit/>

<https://airscope-docxs.readthedocs.io/en/latest/>

The content is well organized and comprehensive. The Readme file is clear.

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Nature-2025-10-27104 - Response to Reviewers' Comments

Response to all Reviewers

We would like to sincerely thank all reviewers for their thorough, thoughtful, and constructive evaluations of our manuscript. All reviewers recognized the significant engineering effort behind Airscope and its potential impact on systems neuroscience, while raising important and valid points that have substantially improved our work. We have addressed every concern rigorously in our revised manuscript through a combination of new experiments, additional analyses, comprehensive language revisions, and expanded open-source materials.

We believe the revision substantively addresses all concerns raised, as detailed below. We summarize these strengthened aspects below, with full point-by-point responses to follow:

1. Independent verification of optical performance. We have completed a rigorous, statistically resolved characterization of resolution and field of view, supported by independent calibration data. We now report (i) a spatially resolved lateral resolution map measured from $n = 6$ beads at each field position ($3.99 \pm 0.08 \mu\text{m}$ at the centre rising to $6.04 \pm 0.55 \mu\text{m}$ at the field edge, 3 mm radius), (ii) full-FOV USAF 1951 and commercial fluorescent-grid calibration images that confirm the 6 mm circular field, and (iii) the axial PSF, now promoted from the Supplementary Information to a main-text panel. All references to resolution have been updated to “4–6 μm across the 6 mm FOV.”

2. Two-photon ground-truth validation of neuron detection. To rigorously verify single-cell counts and signal fidelity at one-photon depth, we constructed a hybrid Airscope/two-photon microscope and acquired 36 temporally interleaved paired recordings across 11 cortical regions in three mice. Across the dorsal cortex Airscope achieves an F1 of 0.78 ± 0.05 , a recall of 0.81 ± 0.08 , and a trace correlation of 0.82 ± 0.14 against the two-photon reference. Layer identity (L2/3) was confirmed by paired DAPI/GCaMP histology. A new table maps every reported neuron count to its specific figure panel.

3. Throughput and temporal-resolution metrics. To rigorously establish the throughput and temporal-resolution metrics, we (i) verified onboard frame rate at $12.47 \pm 0.03 \text{ Hz}$ with $0.17 \pm 0.16 \%$ drop rate from MCU timestamps, (ii) added spatial- and spatiotemporal-bandwidth-product comparisons across all current open-source and commercial wireless miniscopes (revised Supplementary Table 1), and (iii) implemented a higher-speed configuration that achieves 30 Hz at 800×600 by on-sensor binning while remaining fully wireless.

4. Wireless-only enabling experiments and causal controls. To showcase experiments that wireless operation uniquely enables, we performed paired tethered-vs-wireless recordings in the same animals and cortical fields on a 3D terrain arena, showing that tethering changes both vertical movement kinematics and the underlying cortical population geometry compared to Airscope. We also engineered a dual-LED optogenetic Airscope (470 nm imaging + 570 nm optogenetic stimulation) and a microLED-equipped variant used in a simultaneous cortical imaging and seizure-management demonstration in Pentylentetrazol-treated mice.

5. Biological claims, confound controls, and language. To address the pervasive over-interpretation in the original biological text, we have conducted a manuscript-wide revision in which every causal or mechanistic claim has been replaced with observational language (“causal pathway” into “Granger-causal influence”; “unveils cortical basis” into “enables large-scale

population imaging during”). New control analyses explicitly partition motor kinematics (push/retreat/resist), feeding-related behavior, and low-level visual variables from population decoding in the social-hierarchy, social-feeding, and landscape-exploration assays.

6. Hardware refinements during revision. To allow imaging depth to be adjusted after mounting, we have developed and validated a miniature precision-screw mechanism ($\pm 400\ \mu\text{m}$ range, 0.4 g additional weight) integrated into the baseplate. The depth-tuning module is now part of the open-source design package.

7. Open-source release and integrated toolbox. We have substantially expanded our open-source release. The `Airscope_toolbox` repository is now fully public at https://github.com/Neurallabware/Airscope_toolbox and has been organized as an integrated open-source resource. The repository now includes Airscope hardware resources, including Zemax optical files, CAD and 3D mechanical design files, PCB schematics and layouts, and firmware-related resources. It also contains the associated software resources, including acquisition-related code, calcium-imaging analysis pipelines, SAM2Mice, the BERT-based predictor, pretrained checkpoints and example notebooks. In addition, the raw and processed datasets have been organized into a public Airscope data release at https://drive.google.com/drive/folders/1z9ibX8Ob2NnCdJQDHI7Z4tdhdsnpn647B?usp=drive_link. Additional hardware instructions are provided through the Airscope DevKit at <https://airscope.org/devkit/>, and the full online documentation for datasets and analysis workflows is available at <https://airscope-docxs.readthedocs.io/en/latest/>. These updates provide the hardware, software, datasets and documentation needed to facilitate reuse and future development by the community.

8. Figure redesign and manuscript presentation. To present the figures more clearly, we redesigned Fig. 1a to clearly distinguish prior approaches from Airscope, restored the unedited photograph in panel 1m, reordered Fig. 2 so that neuronal imaging is presented first, increased panel resolution to 300 ppi, and added axis labels and colour-bar annotations to every panel previously defined only in the legend.

9. Expanded Discussion. To ground the Discussion more concretely in the broader literature, it now includes a balanced pro/con comparison with head-mounted 2P/3P platforms and incorporates the broader scene-reconstruction prior art (Meyer et al. 2018¹; Parker et al. 2022²; Holmgren et al.³).

We summarize and briefly describe all additional experiments and major revisions in the following table:

Experiment / Revision	Requested by	Description	Related to
Paired tethered vs. untethered recordings	Reviewer 2	Same animals, same cortical fields, paired recordings under tethered and fully wireless configurations on a 3D terrain arena. Tethering distorts vertical movement kinematics and the cortical population geometry that supports them; both are recovered when wireless. PCA trajectories of ascending and descending events are separable before movement	<i>Response Fig. R21, Fig. R22; new Extended Data Fig. 3</i>

Experiment / Revision	Requested by	Description	Related to
		onset; an SVM classifier decodes ascending vs. descending behavior from single trials.	
Optogenetic Airscope prototype	Reviewer 2	Engineered a dual-LED Airscope variant with three 470 nm LEDs for fluorescence imaging and three 570 nm LEDs for optogenetic stimulation, sharing the imaging optical path. Presented as a technological foundation for future causal interrogation rather than a demonstration of cortical causality in this manuscript.	<i>Response Fig. R23</i>
Two-photon ground-truth validation of neuron detection	Reviewers 1, 2, 3	Built a hybrid Airscope/two-photon microscope sharing the sample plane through a 1:1 relay. Performed 36 temporally interleaved paired recordings across 11 cortical regions in three mice. Airscope achieved $F1 = 0.78 \pm 0.05$, $\text{recall} = 0.81 \pm 0.08$, and trace Pearson correlation = 0.82 ± 0.14 against the two-photon reference. Reported as evidence of single-cell signal fidelity and of the reported neuron counts.	<i>Response Fig. R8, R17, R18; Extended Data Fig. 5. Supplementary Fig. 11</i>
Imaging depth and layer identity	Reviewers 2, 3	Confirmed expression in cortical layer 2/3 by paired DAPI staining and two-photon imaging of GCaMP6f in sagittal sections of Rasgrf-Ai148D mice. Used the hybrid 2P/Airscope system to record paired signals at multiple depths beneath the cortical surface, supporting the assignment of imaging depth.	<i>Response Fig. R19 and R20; Extended Data Fig. 5</i>
Spatiotemporal bandwidth product (STBP) comparison	Reviewer 2	Added per-second pixel throughput and spatiotemporal bandwidth product across all surveyed wireless miniaturized microscopes in revised Supplementary Table 1, in addition to the spatial bandwidth product. Reports a $1.67\times$ spatiotemporal throughput lead for Airscope over the closest wired comparator (Kiloscope) once temporal resolution is included, while preserving a $>2\times$ spatial bandwidth product advantage.	<i>Revised Supplementary Table 1; Response Fig. R15</i>
Higher-speed wireless configuration (binning)	Reviewer 2	Applied on-sensor 2×2 binning on the existing OV2640 sensor to raise the frame rate to 30 Hz at 800×600 pixels, matching the temporal resolution of Kiloscope while preserving fully wireless operation.	<i>Response Fig. R28</i>
Depth-tuning mechanism	Reviewer 3	Designed, fabricated, and validated a miniature precision-screw module integrated into the baseplate, providing $\pm 400 \mu\text{m}$ post-implantation focal-plane adjustment and $\sim 0.4 \text{ g}$ additional weight.	<i>Response Fig. R33</i>

Experiment / Revision	Requested by	Description	Related to
Motor-confound analyses for social-hierarchy decoding	Reviewer 3	Quantified neurons tuned to push, retreat and resist; excluded these from win/lose decoders and showed decoding performance was not significantly affected (unpaired t-test; $p = 0.43, 0.37, 0.27$ for accuracy). Population-level motor variables fitted to PCs via GLM explained only a modest fraction of variance, with structured residuals retaining rank information.	<i>Response Fig. R30; Extended Data Fig. 8</i>
Feeding- and visual-confound analyses	Reviewer 3	Extracted feeding-related variables (food-port occupancy, distance to feeder, eating events) and showed that feeding-responsive neurons do not dominate rank decoding in the social-feeding assay. Analogous analyses for the landscape-exploration assay showed that visually responsive neurons do not dominate zone decoding, ruling out visual landmarks as a trivial confound.	<i>Response Fig. R31, R32; Supplementary Fig. 14</i>
Multi-animal simultaneous wireless cortical imaging	Reviewers 1, 2	Simultaneous wireless cortical imaging from multiple freely interacting mice, enabled by Wi-Fi-based bandwidth sharing tested up to five concurrent devices on a single access point with negligible packet loss. Presented as the practical realization of a long-desired experiment, while noting that multi-animal operation is not fundamentally exclusive to Airscope.	<i>Revised main figure; wireless scaling tests in Supplementary Fig. 4</i>
Two-photon validation of axial signal overlap	Reviewers 2, 3	Signals were demixed by the DeepWonder algorithm using a series of morphological and temporal-statistical criteria, and demixing fidelity was further validated by temporally interleaved two-photon and Airscope imaging.	<i>Response Fig. R16; new main-text axial PSF panel</i>
Manuscript-wide language and confound revision	Reviewers 2, 3	Replaced causal/mechanistic claims with observational language throughout the manuscript (e.g., “causal pathway” into “Granger-causal influence”; “unveils cortical basis of ...” into “enables large-scale cortical population imaging during ...”; “initial appraisal of untethered brain-behavior dynamics” into “versatile applications of untethered cortical imaging across diverse behavioral paradigms”). Motor, feeding, and visual confounds explicitly addressed in the relevant assays.	<i>Tracked revisions throughout main text and abstract</i>
Open-source release of full design package	Reviewers 1, 2, 3	Released complete Zemax optical files (.zmx) with lens prescriptions and tolerance/merit-function data; PCB schematics and layout; mechanical CAD (STEP/STL); firmware for onboard MCU and wireless protocol; assembly documentation; and example raw datasets. Itemized inventory added to	<i>Data Availability section; cloud archive</i>

Experiment / Revision	Requested by	Description	Related to
		the Data Availability section with a README in the repository root.	
Unified analysis toolbox and public GitHub	Reviewer 3	Centralized GitHub repository (Airscope_toolbox) consolidating calcium processing pipeline, SAM2Mice, BERT-like predictor, pretrained checkpoints, example datasets and cross-linked documentation as a unified entry point. Added step-by-step tutorials, hardware resources, firmware, analysis notebooks and Airscope DevKit instructions. Repository made fully public and all manuscript links re-verified.	<i>GitHub:</i> https://github.com/Neurallabware/Airscope_toolbox <i>Documentation:</i> https://airscope-docxs.readthedocs.io/en/latest/ <i>Airscope DevKit:</i> https://airscope.org/devkit/
Figure 1, Figure 2 and panel-level redesign	Reviewers 1, 3	Restructured Figure 1 into four thematic rows (concept/throughput; optical design; wireless architecture and assembly; performance benchmarks); redesigned Fig. 1a to clearly distinguish prior approaches from Airscope; restored the unedited photograph in panel 1m; reordered Figure 2 to lead with neuronal fluorescence imaging; raised figure resolution to 300 ppi; replaced parabolic decorative arrows with neutral indicators.	<i>Revised Fig. 1, Fig. 2; numerous panel-level edits</i>
Expanded Discussion and prior-art coverage	Reviewer 2	Rewrote the Discussion to provide a balanced pro/con comparison with head-mounted two-photon and three-photon systems, and added scene-reconstruction prior art (Meyer et al. 2018; Parker et al. 2022; Holmgren et al.).	<i>Revised Discussion; Supplementary Note 5</i>

Taken together, these revisions — twelve newly performed experiments, a manuscript-wide language and confound revision, the complete open-source release of the Airscope_toolbox, a redesigned figure set, and a substantively expanded Discussion — address every blocking and major comment raised by Reviewers 1, 2 and 3. All minor clarifications and panel-level edits requested by the reviewers have likewise been incorporated. We hope this revision conveys the validity, versatility, and expected impact of the Airscope platform. The point-by-point responses are provided below.

Point-by-point response to reviewers' comments

Reviewer #1

Remarks to the Author

Zhang *et al.* present *AirScope*, a ~1-gram wireless miniature mesoscope capable of imaging a 6×6 mm² field of view with approximately 4 μ m lateral resolution and 10 Hz frame rate. The authors demonstrate the system in a range of behavioral contexts, including simultaneous imaging of multiple freely moving mice in room-scale enriched environments, competitive interactions within confined spaces, and single-animal swimming assays. They further present an extension of the platform enabling simultaneous imaging of the brain and spinal cord.

Response: We thank the reviewer for this accurate and concise summary of the manuscript, which captures both the optical specifications and the breadth of behavioral demonstrations. We are particularly grateful that the reviewer recognised the brain–spinal-cord extension as a distinct capability of the platform.

Overall, the manuscript presents an unusually broad and dense body of work, spanning optical instrumentation and multiple neuroscience studies that include claims of novel biological findings. Although the authors adhere to nominal figure count limits, many figures contain a high density of subpanels, with individual figures effectively encompassing complete experimental paradigms. This breadth makes it difficult to rigorously assess all technical and biological claims and raises concern that a standard three-reviewer framework may be insufficient for proper assessment. Ultimately, decisions about scope and structure are editorial, but the editor may wish to consider whether the manuscript would benefit from either an expanded reviewer pool or a more concise presentation achieved by restructuring into more focused submissions, to improve clarity, reviewability, and rigor.

Response: We thank the reviewer for this candid and pragmatic feedback, and we have used it as a central organising principle of the revision. To improve accessibility without sacrificing the integrated nature of the contribution, we have made three concrete changes.

Sharpened biological scope. The neuroscience demonstrations have been restructured around a single flagship experiment that Airscope uniquely enables: simultaneous dorsal-cortex-wide recording from two freely competing mice, allowing us to ask whether cortical activity associated with competitive rank — previously characterised in medial prefrontal cortex⁴ — extends beyond prefrontal circuitry. The formulation tests whether prefrontal findings reflect a broader principle of cortical organization, and it is not addressable with current tools because tethered miniscopes preclude two unrestrained animals from freely encountering each other in tubes, single-port feeders, or running wheels. Secondary biological demonstrations have been consolidated into the Supplementary Information.

Reduced panel density and figure load. Each main-text figure now centres on one experimental claim rather than a complete paradigm. Multi-panel paradigm figures have been trimmed, and methodological subpanels (calibration, validation, control conditions) have been moved into Extended Data and Supplementary Figures, where they remain fully accessible but do not compete for the reader's attention with the primary claim.

Independently assessable raw data and analysis. To allow the reviewer to evaluate technical claims without relying on summary statistics alone, we have released the full Airscope_toolbox repository, including raw recordings, the DeepWonder source-separation network, SAM2MICE behavioral segmentation, and the decoding pipeline. We have also added a paired two-photon ground-truth experiment (36 paired recordings, 1,196 neurons in three mice; $F1 = 0.78 \pm 0.05$, $\text{recall} = 0.81 \pm 0.08$, $\text{trace correlation} = 0.82 \pm 0.14$) as an externally verifiable benchmark of neuronal extraction.

We believe that these changes, together with the detailed point-by-point responses below, will substantially address the reviewer's concern.

With that context, I conducted my review within the scope of my expertise in optical instrumentation and imaging system design. I focused on evaluating whether the instrumentation and performance claims are adequately supported by the presented data. I consider the neuroscience results at a high level, but I am not in a position to assess in depth whether the biological conclusions are fully supported. If the work were to be restructured into more focused submissions, an instrumentation-focused paper could reasonably precede downstream neuroscience applications, and my comments are provided with that perspective in mind.

Response: We are grateful for the reviewer's transparency about the scope of their evaluation, and we have taken the suggestion that an instrumentation-focused paper could reasonably precede neuroscience applications very seriously. After careful consideration, we have chosen to retain a single integrated manuscript, for the following reason: the central claim of the paper is not the optics in isolation but the demonstration that the *combined* engineering of optics, on-board illumination, wireless transmission, and sub-gram form factor enables neuroscience experiments — paired imaging in multiple unrestrained, freely interacting mice. Removing the neuroscience demonstration would leave the engineering contribution without the experimental result that justifies its design constraints.

That said, we have done our best to honour the spirit of the reviewer's suggestion within a single paper. The instrumentation contribution can now be assessed independently of the neuroscience: the optical characterisation has been substantially expanded (full-FOV resolution map, USAF and fluorescent-grid FOV calibration, MTF, transmission bandwidth, head-to-head specification table against the other miniaturized microscopes), and the new two-photon ground-truth experiment provides an externally verifiable benchmark of imaging fidelity at the level of detected neurons and extracted traces. A reader interested only in the device should be able to evaluate the platform on these data alone.

We hope this addresses the reviewer's concern and we appreciate the perspective from which the comments were offered.

Within that scope, this is still an impressive and ambitious body of work, and many of the manuscript's claims are supported by data presented in the Results section. In its current form, the work appears to merit publication in Nature.

Response: We sincerely thank the reviewer for this encouraging assessment.

My primary concerns relate to the interpretation of the field-of-view and resolution characterization data, as well as to several claims that are difficult to independently verify. If these issues are addressed, I would support publication of the manuscript.

Response: We thank the reviewer for stating the path to publication so clearly. The two concerns identified — interpretation of FOV/resolution characterization, and independent verifiability of claims — have driven the principal technical additions of this revision: a multi-bead, statistically resolved resolution map across the full field (Blocking Comment 1a, **Fig. R1**); independent FOV calibration with a commercial fluorescent grid slide (Blocking Comment 2c, **Fig. R7**); a two-photon ground-truth experiment for neuronal detection and trace recovery (Blocking Comment 3, **Fig. R8, Table R1**); and a fully public data and code release (Blocking Comment 4). We address each blocking and non-blocking comment in detail below.

The following comments evaluate whether the stated instrumentation and performance claims are supported by the presented data. Claims are labeled as “blocking” where they affect verification of core system specifications or primary performance conclusions.

Response: We thank the reviewer for adopting this transparent labelling scheme — it has helped us prioritise the revision. Each blocking comment is addressed first within its section, followed by any associated non-blocking points.

Blocking Comments:

Blocking Comment 1 — ~4 μm lateral resolution

1a — Statistical characterization of resolution across FOV

At a minimum, this claim should be slightly tempered. Supplemental Fig. S10 shows bead-derived lateral widths increasing from approximately 3.9 μm to 4.8 μm across the field of view, indicating measurable variation with field position. While the example data are useful, the analysis would benefit from a more quantitative and statistically representative treatment. For example, lateral FWHM measurements from multiple beads (e.g., 4–6) near the center of the field could be averaged and reported with standard deviation, and the same analysis repeated at multiple positions toward the edges of the field. This would help clarify whether the presented beads are representative and how resolution degrades at field extremes, where performance is expected to be worst.

Response: We thank the reviewer for this constructive suggestion. In this revision, we have performed a comprehensive, statistically rigorous resolution characterization:

1. Multi-bead FWHM analysis with statistics. Following the reviewer's suggestion, we measured lateral FWHM from multiple beads ($n = 6$) at each field position, reporting mean $\pm$ standard deviation. At the field center, the lateral FWHM is $3.99 \pm 0.08 \mu\text{m}$; at the field edge (3 mm from the center), it increases to $6.04 \pm 0.55 \mu\text{m}$, as shown in **Fig. R1**. Meanwhile, we verified that in such a resolution range, the neuronal signals can still be very well verified through the paired two-photon microscope recordings, in **Fig. R17**.

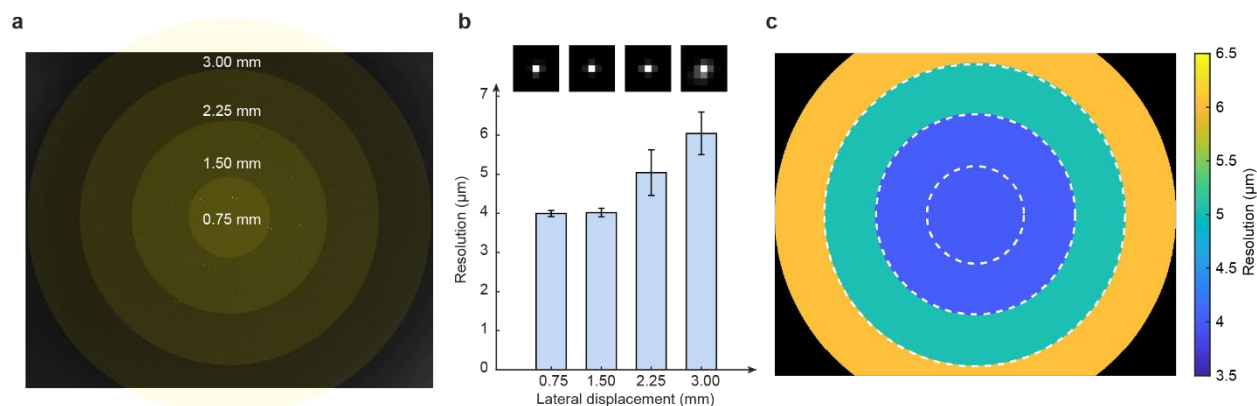


Fig. R1. Resolution characteristics of Airscope across the field of view (FOV). **a**, Imaging of $1\ \mu\text{m}$ point emitters across the FOV. Colored concentric circles indicate regions at increasing radial distances (0–3 mm) from the center of the FOV. **b**, Quantification of lateral resolution as a function of lateral displacement from the FOV center. Resolution was measured from imaged point emitters and averaged over two orthogonal directions ($n = 6$ per region, mean $\pm$ S.D.). Representative examples of the imaged spots are shown above the bar plot. **c**, Spatial distribution of lateral resolution across the field of view, each region is color-coded according to the mean resolution value.

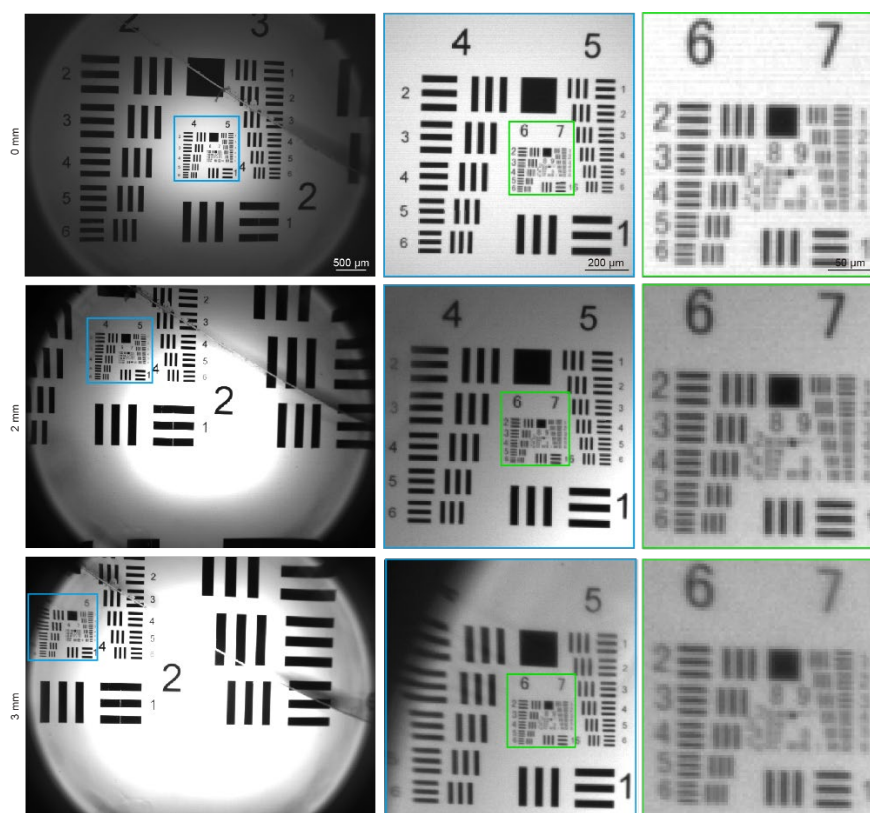


Fig. R2. Airscope resolution validation using USAF 1951 resolution target at different positions. Zoomed-in views at second and third columns.

2. FOV-aware resolution map. We have performed resolution measurements across four radially equidistant annular regions spanning the full field of view, creating a spatially resolved resolution map in **Fig. R1** (also in revised **Supplementary Fig. 8**). This map clearly shows the expected gradual degradation from center to periphery, consistent with the optical design specifications.

3. Resolution validation using USAF 1951 resolution target. To further support the resolution characteristic results, we also imaged a standard USAF 1951 resolution target at different positions across the FOV. At the center of the FOV, Group 7 Element 1 (line width of 3.91 μm), was clearly discernible. Toward the edge of the FOV, Group 6 Element 4 (line width of 5.52 μm) even Group 6 Element 5 (line width of 4.92 μm), remained discernible (**Fig. R2**).

4. Revised resolution claim. In the original manuscript, the reported central and peripheral resolution values, approximately 3.9 μm and 4.8 μm , respectively, were obtained from representative beads measurements rather than from a systematic statistical characterization across the full FOV. In addition, the previous “peripheral” measurement did not sample the outermost region of the 6 mm FOV. On the basis of the more comprehensive analysis presented above, we have revised the resolution claim throughout the manuscript to 4–6 μm lateral resolution across the 6 mm FOV. Supplementary Table 1 has also been updated to include both central and peripheral resolution values. In addition, the space–bandwidth product calculation has been revised accordingly using this updated resolution range. This revised description more accurately reflects the spatial variation of the optical resolution across the full FOV.

1b — Pixel sampling and bead profile processing

In addition, given the stated ~ 6 mm field of view and a ~ 2 MP image sensor, the implied sampling at the specimen plane is on the order of ~ 3 –4 μm per pixel, depending on the sensor format. At this sampling density, a ~ 4 μm feature would be expected to span approximately one pixel, and raw bead intensity profiles would therefore appear quantized rather than smooth. The bead profiles shown in Supplemental Fig. S10 appear smoother than would be expected from raw pixel sampling at this scale. Please clarify the effective pixel size at the specimen plane and whether the bead images or intensity profiles were rescaled, interpolated/upsampled, averaged across frames, filtered, deconvolved, or fit to a model, and whether the plotted curves represent raw pixel samples or fitted profiles.

Response: We thank the reviewer for this careful technical observation, and we sincerely apologize for the lack of clarity in the original figure. The effective pixel size at the specimen plane in our system is approximately 3.7 $\mu\text{m}/\text{pixel}$. At this sampling density, we agree that a ~ 4 μm feature is expected to span approximately one pixel in the raw data.

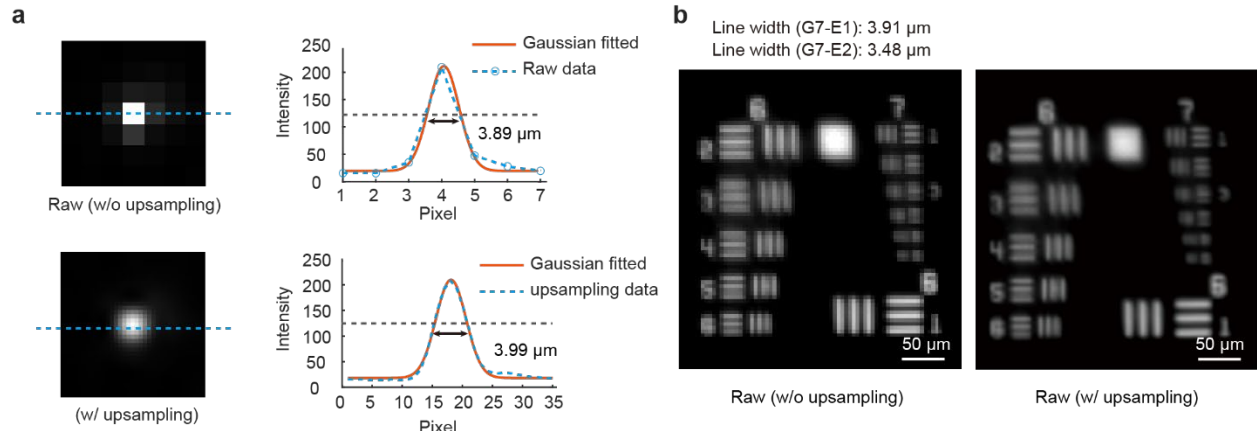


Fig. R3. Comparison of raw and upsampling processing over point emitters and a USAF 1951 resolution target. **a**, Left panel: raw (without upsampling) and $5\times$ upsampled images of a single point emitter. Right panel: corresponding intensity profiles along the dashed line, showing both the original data and Gaussian fitting curves. **b**, Raw (top) and $5\times$ upsampled (bottom) images of the USAF 1951 resolution target.

For visualization purposes only, the bead images shown in original Supplemental Fig. 10 were upsampled by a factor of 5 to improve visual clarity. Importantly, no deconvolution, filtering, or other image enhancement procedures were applied. To avoid any ambiguity, we have now included the original, non-upsampled bead images and corresponding intensity profiles in the revised supplementary materials (**Fig. R3a**). As expected, these raw data show that the bead signal is primarily confined to approximately a single pixel, which is consistent with the stated sampling and supports our $\sim 4 \mu\text{m}$ resolution claim.

For quantitative resolution estimation, we extracted intensity profiles and performed Gaussian fitting to measure the full width at half maximum (FWHM). The measured FWHM is approximately 1.1 pixels, corresponding to $\sim 4 \mu\text{m}$ in the specimen plane. We emphasize that this analysis was performed on the raw (non-upsampled) data. Applying the same fitting procedure to the upsampled data yields consistent FWHM values (3.99 μm), confirming that upsampling does not affect the resolution measurement and serves only a visualization purpose.

In addition, to mitigate potential ambiguity associated with bead-based measurements, we have included both raw and upsampled images of a USAF 1951 resolution target. The system is able to resolve Group 7, Element 1 even Element 2, further supporting a spatial resolution on the order of $\sim 4 \mu\text{m}$ (**Fig. R3b**).

To further validate the optical resolution of Airscope, we performed an additional measurement using a high-resolution Sony 200 MP sensor. After calibration, the corresponding object-plane sampling was 1.28 μm per pixel. Using the Airscope lens module coupled to this high-resolution sensor, we imaged 1 μm point emitters at the center of the FOV without upsampling or deconvolution. The imaged point emitter occupied approximately 3 pixels, corresponding to ~ 3.7 – $3.8 \mu\text{m}$ in the object plane, which is consistent with the claimed $\sim 4 \mu\text{m}$ central resolution (**Fig. R4a, b**). This result confirms that the reported resolution is not an overclaim caused by the original measurement. To further support this conclusion, we also imaged a standard USAF 1951 resolution target at different positions across the FOV. At the center of the FOV, Group 7 Element 3 (line

width of $3.10\ \mu\text{m}$), was clearly discernible. Toward the periphery, Group 6 Element 4 (line width of $5.52\ \mu\text{m}$) even Group 6 Element 5 (line width of $4.92\ \mu\text{m}$), remained discernible (**Fig. R4c**). Together, these measurements independently support a lateral resolution of approximately $4\text{--}6\ \mu\text{m}$ across the $6\ \text{mm}$ FOV.

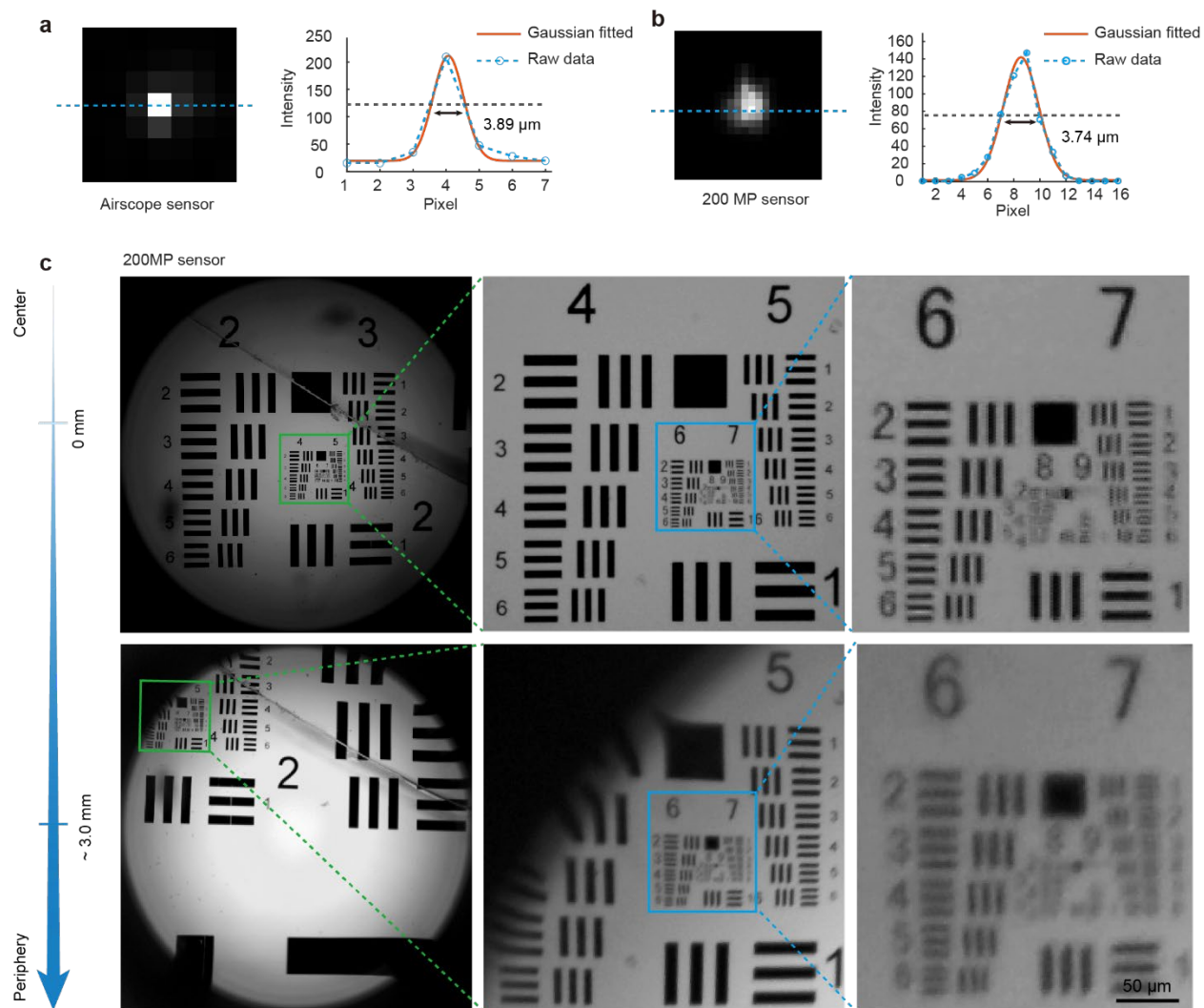


Fig. R4. Airscope resolution validated using a high-resolution 200 MP sensor with point emitters and a USAF 1951 resolution target. **a**, Image of a $1\ \mu\text{m}$ point emitter acquired with the Airscope sensor, with the corresponding intensity profile measured along the dashed line. Raw data are shown in blue, and the Gaussian fit is shown in orange. **b**, Image of the same $1\ \mu\text{m}$ point emitter acquired using the Airscope lens coupled to a high-resolution 200 MP sensor. The corresponding intensity profile is shown with the same color as in **a**. **c**, USAF 1951 resolution target images acquired using the 200 MP sensor at different positions across the FOV. Zoomed-in views at second and third columns. Scale bar: $50\ \mu\text{m}$.

We have supplied the above new characterization into **Supplementary Fig. 8**.

Blocking Comment 2 — 6 mm diameter field of view

2a — Internal inconsistency in Supplemental Fig. S10A

There appears to be an internal inconsistency between the stated field of view and the spatial scaling in Supplemental Fig. S10A. If the scale bar is taken at face value, a 6 mm field of view is achieved; however, under that same scaling, the beads appear to be approximately 18–20 μm in diameter rather than 4 μm . Conversely, if the beads are assumed to be 4 μm , the implied scale bar would correspond to $\sim 120 \mu\text{m}$, yielding a field of view closer to $\sim 1.5 \text{ mm}$. These interpretations are not mutually consistent.

Response: We thank the reviewer for this meticulous analysis, and we sincerely apologize for the confusion. We respectfully but firmly maintain that there is no inconsistency in the underlying measurement: the image covers approximately 6 mm, while the system resolution remains on the order of 4 μm .

The source of the confusion is that the image originally shown in **Supplementary Fig. 10a** was intended only as a visualization of point imaging across the full field, rather than as a quantitative size reference. Because a 4 μm bead occupies only a tiny fraction of a 6 mm-wide image ($4 \mu\text{m} / 6 \text{ mm} \approx 1/1500$, i.e., $\sim 0.067\%$ of the full image width), the raw signal is barely visible when displayed over the full field of view. To make the point signals visible across the entire image, we applied strong contrast stretching. This causes the bright spots to exceed the display dynamic range and appear saturated, making them look substantially larger than their true physical size.

To further clarify this point, we added a comparison figure acquired from a uniformly spaced dot array. In this comparison, panel **a** of **Fig. R5** shows the over-stretched rendering used for visualization, whereas panel **b** of **Fig. R5** shows the same data displayed at a normal contrast level (both images were upsampled by 3x for easier viewing). This comparison makes clear that the enlarged spot size in **a** is a display pattern caused by visualization settings, rather than a reflection of the true bead size or the system resolution. The resolution estimate thereby should not be based on the apparent spot diameter in a contrast-enhanced full-field rendering but through rigorous analysis, which we have supplied in this review, **Fig. R1** and **Fig. R3**. Similarly, the FOV should be independently verified, see our new verification in **Fig. R6**.

We agree with the reviewer that the original presentation may be misleading, even though it was intended only as a schematic illustration to show that point-like objects were imaged over the full field. Thereby, we have replaced the original **Supplementary Fig. 10** (now **Supplementary Fig. 8**) with an unprocessed representation of the zoom-in view of the point emitters and the FOV-aware resolution map.

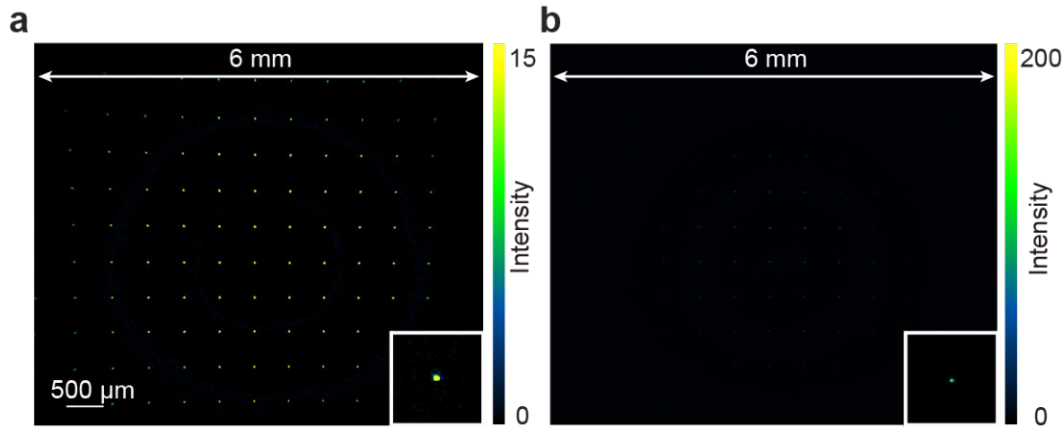


Fig. R5. Comparison of over-stretched and properly adjusted contrast in imaging of a 1 μm point emitter array. **a**, Image with over-stretched contrast, leading to intensity saturation. **b**, Image with properly adjusted contrast. Insets show magnified views of representative emitters.

2b — USAF target scaling discrepancy

A similar discrepancy arises when using the Air Force test target for calibration: Group 5, Element 1 should have a line width of approximately 15 μm , which—under the apparent scaling—corresponds to a field of view of roughly 1.3 mm, consistent with the bead-based estimate above. Given these inconsistencies, clarification is needed.

Response: We thank the reviewer for this thorough cross-validation. The USAF resolution target image in the original submission showed only a cropped sub-region of the full field of view. Group 5, Element 1 (line width $\sim 15.6 \mu\text{m}$) was used for local resolution assessment within that sub-region, not for FOV calibration.

In this revision, the full-FOV USAF target image is now presented (**Fig. R6**), in which all elements of Group 2 and Group 3 are fully contained within the field of view. The measured distance between known features at opposite ends of the field confirms a diagonal extent of greater than 6 mm, fully consistent with our stated 6 mm circular FOV. A zoomed-in view is also shown in the right panel. All scale bars have been verified against multiple independent target features.

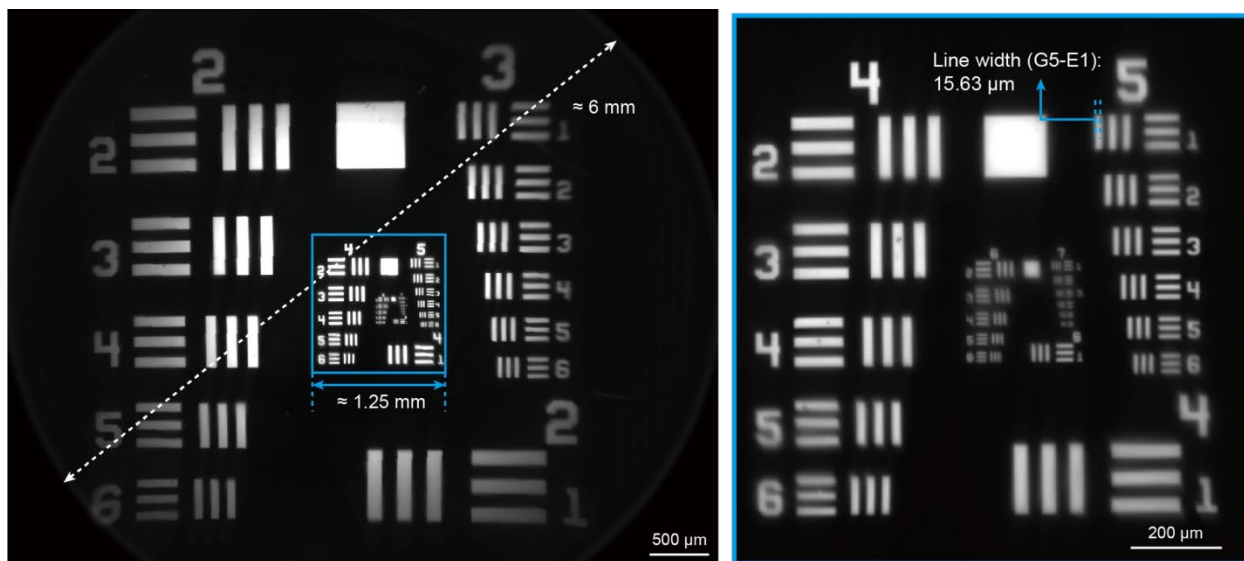


Fig. R6. Full field-of-view imaging of a USAF resolution target, with a zoomed-in view shown on the right panel.

To avoid ambiguity, we have added the full field-of-view image of the resolution target in the revised manuscript (see new **Supplementary Fig. 8**).

2c — Independent calibration with fluorescent grid slide

Providing raw calibration data or an image of a fluorescent grid slide would help definitively establish the field of view. Commercial fluorescent grid targets have historically been available (e.g., via Max Levy), and may still be obtainable through the manufacturer. Independent verification of this calibration would require either unambiguous calibration measurements (e.g., imaging of a fluorescent grid slide) or access to the underlying raw calibration data, neither of which appears to be included in the current data repository.

Response: We thank the reviewer for this suggestion. We have followed the reviewer's recommendation and acquired a calibration image using a commercial fluorescent grid slide (grid spacing: 0.1 mm). The resulting image clearly shows the grid pattern spanning the full sensor field, and counting grid lines across the field diameter yields an FOV measurement about 6.1 mm (**Fig. R7**), in excellent agreement with both the USAF target measurement and the optical design specification (**Fig. R6** and new **Supplementary Fig. 8**).

This fluorescent grid calibration image and USAF resolution target images have been added to the open-source data repository to enable fully independent verification of the field-of-view specification.

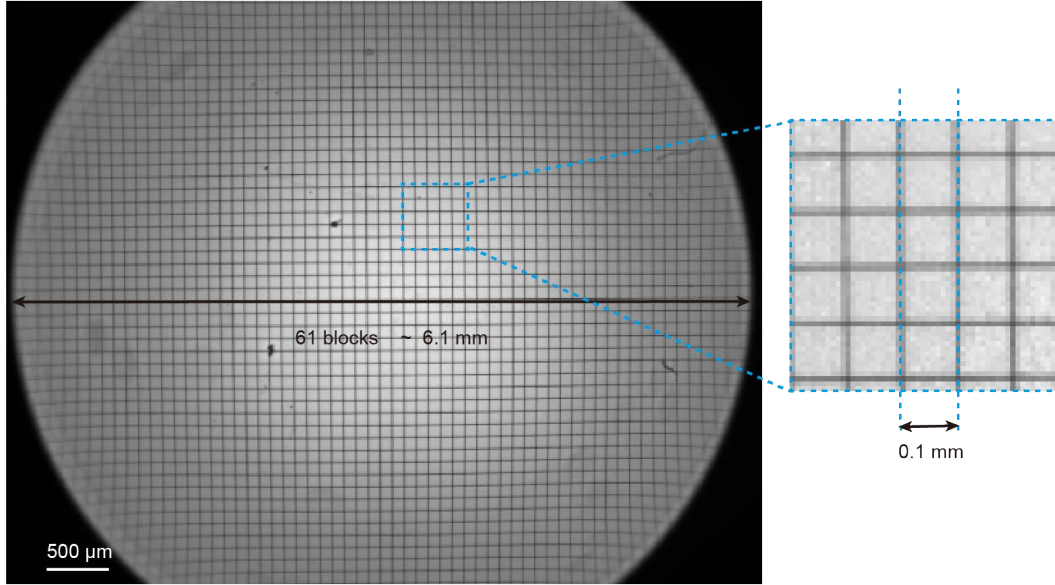


Fig. R7. Field of view measurement using commercial fluorescent grid targets with a grid spacing of 0.1 mm.

Blocking Comment 3 — Recording from >8,500 neurons per animal (~25-fold increase)

Supplementary Table 2 indicates that this claim is supported by Fig. 2; however, neuron counts are not explicitly quantified in the main text, figure captions, or accompanying analysis. While panels c, g, and i visually suggest large neuronal populations, the actual counts cannot be independently verified from the presented figures alone. Substantiating this claim would require either explicit quantitative summaries derived from the underlying segmentation results (e.g., reported neuron counts per animal with clearly defined inclusion criteria) or access to the underlying data and analysis pipeline. As currently presented, this claim is not independently verifiable.

Response: We thank the reviewer for this important concern regarding verifiability. We fully agree that independently verifiable neuron counts are essential. In this revision, we have addressed this at three levels:

1. Explicit quantitative summaries. We have created a comprehensive table (**Table R1**) that reports neuron counts for every experimental panel. This table directly maps each reported number to the corresponding figure panel.

2. Open-source analysis pipeline with clearly defined inclusion criteria. The complete DeepWonder software with all parameter configurations used in this study has been released as part of the Airscope_toolbox repository. Inclusion criteria are defined in the configuration files, and we provide a step-by-step tutorial demonstrating how to reproduce the neuron detection results from the raw data provided in the repository. Importantly, the same inclusion criteria were applied uniformly across all datasets:

- Neuronal diameter $\geq 10 \mu\text{m}$.
- Elliptical roundness (major/minor axis) ≤ 3 .

- Mask occupancy $\geq 50\%$, where occupancy = (mask pixels inside bounding box) / (bounding-box width $\times$ bounding-box height)
- Temporal consistency across ≥ 5 consecutive frames
- Probability-map confidence through DeepWonder network ≥ 0.5
- Centroid not located within the vessel or outside-FOV exclusion zone

These criteria, together with the DeepWonder neuronal segmentation procedures, are reported in full in new **Supplementary Note 2**.

3. Two-photon imaging provides ground-truth validation supporting the fidelity of neuronal recording and extraction.

To validate neuron detection performance, we performed temporally interleaved two-photon and Airscope recordings over the same cortical regions in the same animals. A hybrid imaging system was constructed, in which a 1:1 relay optical path was introduced to couple the sample to the Airscope module, while a two-photon excitation/detection arm (920 nm) was integrated for ground-truth imaging. A 470-nm LED (M470L5, Thorlabs) was used for single-photon excitation in the Airscope path (as in **Fig. R17**). For data analysis, neuronal activity was extracted from Airscope recordings using the same processing pipeline. For the two-photon data, extraction was performed using Suite2p followed by semi-manual curation to establish ground truth.

Across three mice, we acquired 36 paired recordings spanning the dorsal cortex, each recording covering approximately $0.35 \text{ mm} \times 0.3 \text{ mm}$ ($\sim 0.1 \text{ mm}^2$). The Airscope recording achieved an F1 score of 0.78 ± 0.05 , a recall of 0.81 ± 0.08 , and a temporal trace activity correlation of 0.82 ± 0.14 (mean $\pm$ SD, $n = 1,196$ neurons) relative to the reference recorded by two-photon microscope, as in **Fig. R8**. More detailed statistical results are provided in **Fig. R17** and **R18**.

These results provide a validation of the neuron detection and extraction performance and support the reported neuron counts.

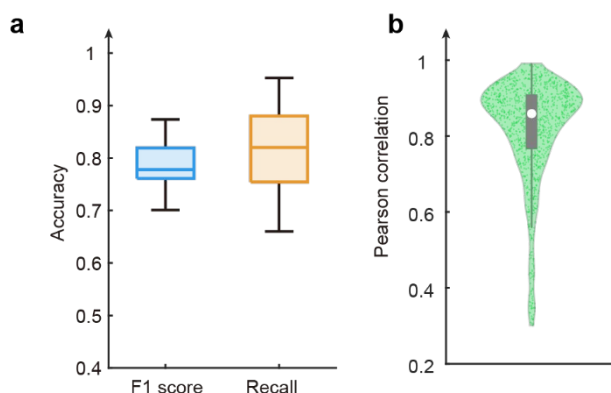


Fig. R8. Quantitative validation of Airscope neuron detection performance using temporally interleaved two-photon recordings as reference. **a**, F1 score and recall of neuron detection by Airscope compared with two-photon ground truth data presented by boxplots across 3 animals and 36 recording sessions. (Central line inside the box: median. Box: interquartile range. Whiskers: maximum and minimum). **b**, Pearson correlation score of temporal activity of neuron between

Airscope and two-photon recordings. (n = 1,196 neurons, white circle: median. Thick gray vertical line: interquartile range. Thin vertical lines: upper and lower proximal values. Transparent disks: data points. Transparent violin-shaped areas: kernel density estimate of data distribution).

Table R1. Neuron counts for every experimental panel

Animal ID	Figure panel	Verified neurons
14_#autism	Fig. 2k	8583
8#	Fig. 2l-n, Supplementary Fig. 9d-h	4560
1#_spine	Fig. 6e-g	4253
3_new#	Fig. 3d, Fig. 3h, Fig. 3i, Extended Data Fig. 7d-e,	5,944
7#	Fig. 3g, Fig. 5e-f, Fig. 5h-i, Extended Data Fig. 9h, Extended Data Fig. 9m, Extended Data Fig. 10 c-d, Extended Data Fig. 10 f-g, Extended Data Fig. 10 k-l, Extended Data Fig. 7d-e, Supplementary Fig. 15a-b, Supplementary Fig. 15d-e,	7133
20#	Fig. 3i, Fig. 5e-f,	5,823
5_new#	Fig. 3i, Extended Data Fig. 7d-e,	7,709
swim#	Extended Data Fig. 9a-g,	2,657
16#	Fig. 5e-f, Fig. 5i, Extended Data Fig. 9h, Extended Data Fig. 9k-m, Supplementary Fig. 15d-e,	3778
6#	Extended Data Fig. 9h, Extended Data Fig. 9m,	2,643
80#	Extended Data Fig. 9h, Extended Data Fig. 9m, Extended Data Fig. 7d-e,	1,736
6_new#	Extended Data Fig. 7d-e,	5,242
1#_cage_3	Fig. 4b, Fig. 4f-m, Fig. 4p-r, Extended Data Fig. 8, Supplementary Fig. 14,	5,609
2#_cage_3	Fig. 4b, Fig. 4f-m, Fig. 4p-r, Extended Data Fig. 8, Supplementary Fig. 14,	5,331
3#_cage_3	Fig. 4b, Fig. 4f-m, Extended Data Fig. 8,	3714
4#_cage_3	Fig. 4b, Fig. 4f-m, Extended Data Fig. 8,	3,893
1#_cage_2	Fig. 4b, Fig. 4f-m, Fig. 4p-r, Extended Data Fig. 8,	4,574
2#_cage_2	Fig. 4b, Fig. 4f-m, Extended Data Fig. 8,	1,834
3#_cage_2	Fig. 4b, Fig. 4f-m, Fig. 4p-r, Extended Data Fig. 8, Supplementary Fig. 14,	6,755

4#_cage_2	Fig. 4b, Fig. 4f-m, Extended Data Fig. 8, Supplementary Fig. 14,	4,962
1#_cage_1	Fig. 4b, Fig. 4c-e, Supplementary Fig. 14,	3812
2#_cage_1	Fig. 4b,	4,573
3#_cage_1	Fig. 4b,	3,326
4#_cage_1	Fig. 4b, Fig. 4c-e, Supplementary Fig. 14,	5807
1#_3D_terrain	Extended Data Fig. 3u-v	3921
2#_3D_terrain	Extended Data Fig. 3u-v	3944
3#_3D_terrain	Extended Data Fig. 3u-v	3143

Blocking Comment 4 — Data and Code Availability

The manuscript states that raw data and analysis code are publicly available; however, the GitHub repository listed in the manuscript currently returns a 404 error and is not accessible. In addition, the referenced cloud storage folder contains only behavioral data (e.g., video recordings of animal movement) and does not appear to include any raw data generated by the AirScope device itself (e.g., sensor-level recordings, calibration data, or neuronal activity traces prior to processing). The absence of underlying raw device data precludes independent verification of several core system specifications and performance claims, including frame rate, optical resolution and field-of-view calibration, and reported neuron counts. Because the manuscript explicitly asserts public availability of raw data, resolving this discrepancy is necessary for verification, reproducibility, and evaluation of the reported claims. The editor should confirm that a valid, accessible archive containing the underlying raw device data and analysis code exists prior to publication.

Response: We sincerely apologize for this oversight. The GitHub repository was inadvertently left in private mode during the submission process, and the cloud storage folder was incompletely populated. We recognize that this discrepancy between the manuscript's data availability claims and the actual state of the repositories is unacceptable, and we have taken comprehensive corrective action.

1. GitHub repository and online resources. The repository is now public and fully accessible at https://github.com/Neurallabware/Airscope_toolbox. It contains the complete Airscope codebase, including firmware and acquisition-related resources, calcium-imaging analysis pipeline, SAM2Mice for multi-animal segmentation, the BERT-based social competition predictor, pretrained model checkpoints, and analysis scripts used in this study. Installation instructions, dependency lists, example datasets, and Jupyter notebooks for loading and visualization are included.

In addition, the following online resources are now publicly available:

- <https://airscope.org/devkit/> — hardware assembly, firmware flashing, acquisition setup, and platform-extension instructions.
- <https://airscope-docs.readthedocs.io/en/latest/> — complete documentation for Airscope hardware, software, datasets, and analysis workflows, including usage examples for the calcium-imaging pipeline, SAM2Mice, and the BERT-based predictor.

2. Raw and processed experimental data. The public data release is now fully populated and accessible at https://drive.google.com/drive/folders/1z9ibX8Ob2NnCdjqDHI7Z4tdhdspn647B?usp=drive_link. The release includes:

- **Raw calcium-imaging recordings:** MJPEG frames as captured by the onboard sensor. The data release additionally provides step-by-step Jupyter notebooks for calcium-imaging processing and signal extraction in https://github.com/Neurallabware/Airscope_toolbox/tree/main/Software/Airscope_ca_processing/notebooks
- **Multi-animal interaction dataset:** Including extracted calcium traces, corresponding behavioral videos, SAM2Mice segmentation masks, neuronal masks, and brain-region parcellations. Example Jupyter notebooks for data loading and visualization are also provided in https://github.com/Neurallabware/Airscope_toolbox/tree/main/Data_release/experiments/multi_animal_interaction.
- **Enriched-habitat dataset:** Including extracted calcium traces, synchronized multi-view behavioral videos, tracking bounding box, neuronal masks, and brain-region parcellations. Example Jupyter notebooks for data loading and visualization are also provided in https://github.com/Neurallabware/Airscope_toolbox/blob/main/Data_release/experiments/enriched_habitat/.
- **SAM2Mice benchmark data:** Including example videos for segmentation and annotated ground-truth for MOTA calculation. Example Jupyter notebooks to calculate MOTA compared to DeepLabCut and SuperAnimal is provided in https://github.com/Neurallabware/Airscope_toolbox/blob/main/Software/SAM2Mice/SAM2_Mice/benchmark/.
- **Optical calibration data:** USAF 1951 target images, and fluorescent grid slide raw images, which are provided in https://github.com/Neurallabware/Airscope_toolbox/tree/main/Data_release/optical_characteristic/.
- **MCU timestamp data:** Frame timing data for independent frame rate verification.

3. Expanded Data Availability and Code Availability statements. We have comprehensively revised both sections of the manuscript to provide a detailed itemized inventory of all shared resources, including:

- **Optical design:** Zemax files (.zmx) with full lens prescriptions and optimization merit functions.
- **Electronics:** PCB schematics and layout files, bill of materials.
- **Mechanical:** CAD models for all housing, baseplate, and depth-tuning components (STEP and STL formats).
- **Firmware:** Source code for onboard MCU and wireless transmission protocol.
- **Software:** Full analysis pipeline including calcium-imaging analysis pipeline, SAM2Mice, and BERT-based predictor.
- **Documentation resources:** Assembly instructions, alignment procedures, firmware flashing procedures, acquisition-software setup, and platform-extension documentation is available at <https://airscope.org/devkit/>. Full online documentation for the Airscope

hardware, software, datasets, and analysis workflows is available at <https://airscope-docxs.readthedocs.io/en/latest/>.

- **Data resources:** Including raw calcium-imaging recordings, multi-animal interaction and enriched-habitat datasets, SAM2Mice benchmark data, optical calibration data, and MCU timestamp data, together with example Jupyter notebooks for data loading, visualization, processing, and benchmarking.

A README file in the repository root directory provides the complete inventory with direct links. We have verified that all links are functional and all files are accessible.

Non-Blocking Comments

Non-Blocking Comment 1 — 10 Hz frame rate verification

The manuscript states multiple times that the system operates at a 10 Hz frame rate; however, it is not clear how this value can be independently verified from the presented data. The authors should clarify how the effective frame rate was measured (e.g., acquisition timing, dropped frames, or representative timestamps), either in the Methods or Supplementary Information. Independent verification of this claim would also require access to the underlying raw device data, which do not appear to be included in the current data repository.

Response: We thank the reviewer for this concern regarding frame rate verification. During revision, we have updated the firmware; now the Airscope support full frame output at 12.5 Hz. In the following, we provide comprehensive verifications over the frame rate. Metrics of the following are calculated as follows:

- **Frames per second (FPS):** Defined as the reciprocal of the mean inter-frame interval. The inter-frame interval is computed as the difference between consecutive MCU hardware timestamps.
- **Drop rate:** Defined as the fraction of expected frames that were not successfully acquired or stored during a recording session. We defined a frame drop event as occurring when two frames have interval larger than 160 ms. Calculated as (number of missing or corrupted frames) / (total expected frames) $\times$ 100%.

Based on these results, as illustrated in **Fig. R9**, we confirmed that Airscope maintains a frame rate of 12.47 ± 0.03 Hz. The frame loss rate remains negligible at $0.17\% \pm 0.16\%$. To enable independent verification, the raw MCU hardware timestamps can be downloaded from https://drive.google.com/drive/folders/1EPN6YSidfCAPXQBSYq9nOefl76L-_FHE, and the corresponding analysis code for frame-rate calculations is provided on GitHub at https://github.com/Neurallabware/Airscope_toolbox/tree/main/Data_release/MCU_timestamp.

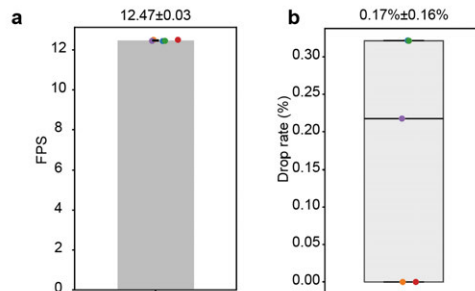


Fig. R9. Airscope frame rate verification. a. Frame rate. b. Drop rate. Values are presented as mean $\pm$ SD from five repeated tests on the same device.

Non-Blocking Comment 2 — ~1 g total device mass

The reported ~1 g device mass is an important enabling parameter, but the basis for this value is not explicitly documented. While Fig. 1 (panels n and o) qualitatively suggests minimal impact on animal motion, it is unclear what components are included in the stated mass (e.g., optics, housing, battery, encapsulation). Explicitly stating how the mass was measured and what configuration was used would strengthen this claim.

Response: We thank the reviewer for this suggestion. The baseline device consists of six primary components, including the housing, an LED-integrated FPC, PCB, emission filter, 4 excitation filters, and lens, which are encapsulated using photosensitive adhesive, resulting in a total mass of 1.0093 g after encapsulation. We have added a detailed component-level mass breakdown in a new figure panel (**Fig. R10**) and an accompanying table (**Table R2**), as following.

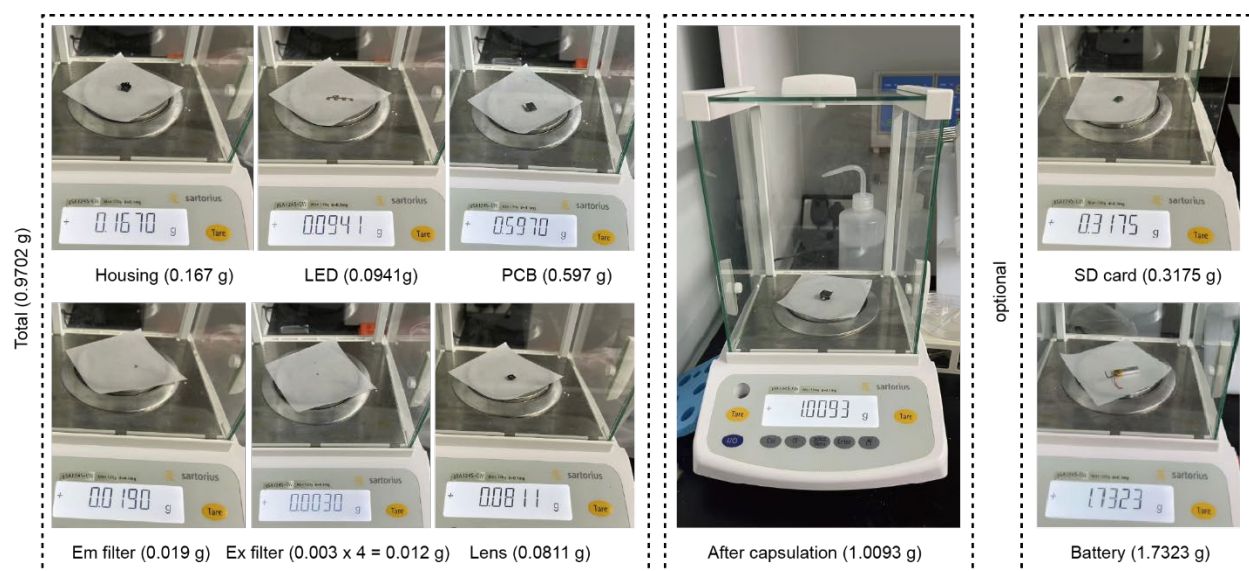


Fig. R10. Weight measurement of each part and fully assembled device of Airscope.

Table R2. Component weight of Airscope

Component	Measured parts	Mass (g)
Mechanical part	Housing	0.1670
Optical module (lenses + filters)	Lens + Em filter + Ex filter $\times 4$	0.1121
Image sensor + PCB	PCB	0.5970
LED illumination module	LED	0.0941
Total (microscope only)		0.9702
Encapsulation	Glue	<0.05

Total (after encapsulation)		1.0093
SD card (optional)	SD card	0.3175
Battery (optional)	Battery (1.7 g config)	1.7323

The stated "~1 g" refers to the microscope body without SD card and battery, as these are modular and interchangeable. The mass was measured using an analytical balance (resolution: 0.0001 g).

The above updated specifications are now included in the revised manuscript in **Supplementary Table 4**.

Non-Blocking Comment 3 — ~2× reduction in optical track length

This claim appears in the Abstract but is not clearly developed elsewhere in the manuscript. Aside from a brief mention in the Fig. 1 caption, I do not see a quantitative comparison that supports a twofold reduction relative to prior systems. Supplemental Fig. S2 illustrates how the custom aspheric lenses contribute to a shortened conjugate, but the claim would benefit from being restated more explicitly and supported within the context of the authors' own design and measurements.

Response: We thank the reviewer for identifying this gap between the Abstract claim and its support in the main text.

In this revision, we provide a rigorous, quantitative comparison of optical track lengths (defined as physical distance from the anterior surface of the primary lens element to the final image), as shown in **Fig. R11**. By evaluating the Zemax files from the respective open-source repositories, representative spherical systems, such as Miniscope V4 and MiniXL, exhibit total track lengths of 22.44 mm and 22.88 mm, respectively, indicating that Airscope's total track length is more than six-fold shorter (3.36 mm of Airscope). Other models, including Miniscope V3 and MiniLFOV, feature even greater track lengths.

Furthermore, compared to the aspherical lens system of Kiloscope, Airscope's total track length is 2.84 times shorter, despite Airscope's superior FOV and a longer working distance optimized for onboard illumination implementation. Detailed comparisons between the optical architectures of Airscope and Kiloscope are provided in our response to Reviewer 2, Major Comment 1.

To substantiate these claims, we have revised the Abstract as follows: "*An aspheric-lens design achieves a 2.84× shorter total track length than previous miniaturized microscopes while matching the neuron yield of benchtop one-photon mesoscopes.*" We further added the **Fig. R11** as support evidence into **Supplementary Fig. 1**.

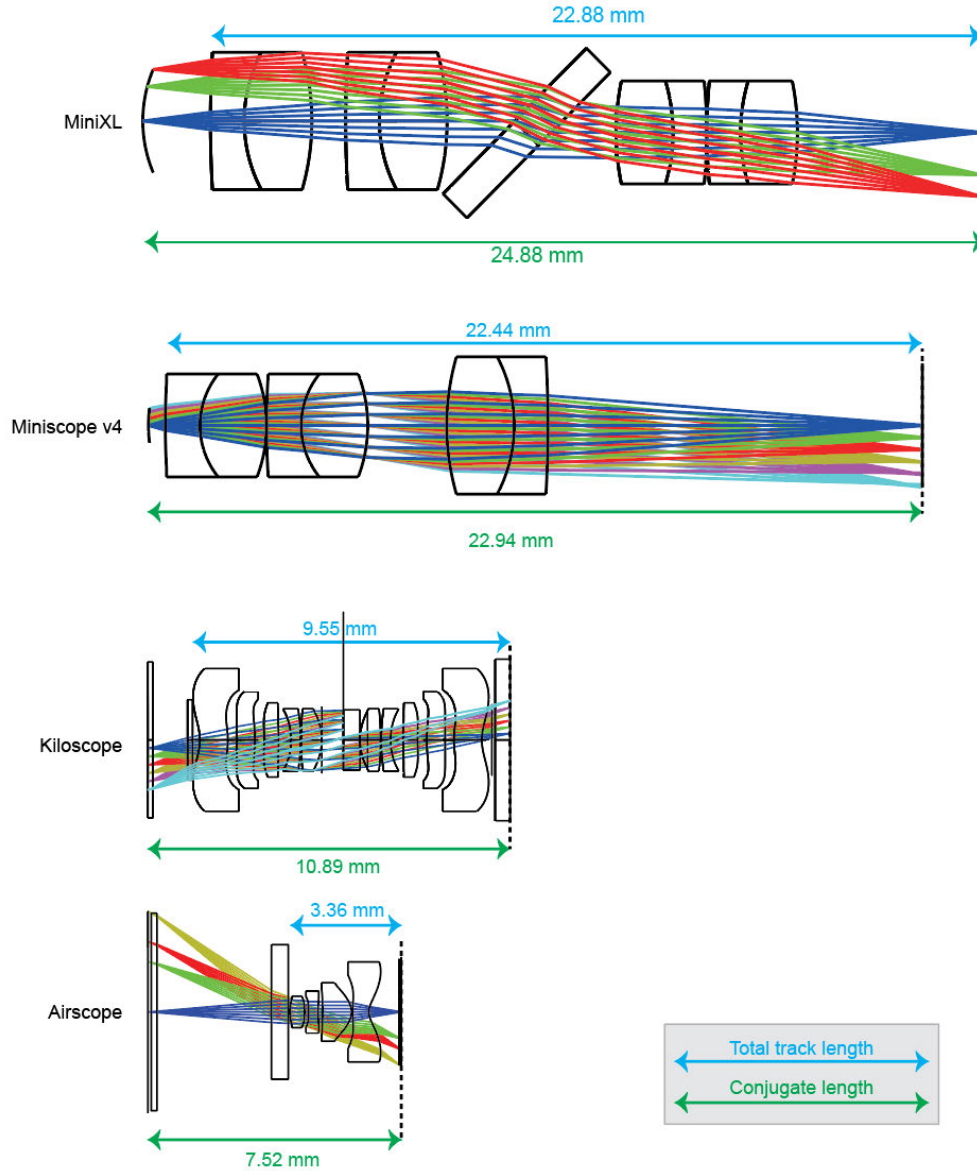


Fig. R11. Total track length comparison among Miniscope V4, MiniXL, Kiloscope, and Airscope.

Non-Blocking Comment 4 — Cortical imaging under "previously inaccessible" conditions

The manuscript states that AirScope enables cortical imaging under conditions that were “previously inaccessible”. Several specific examples are given, but the basis for exclusivity is not always clear.

Response: We thank the reviewer for this careful examination of our claims. We acknowledge that "previously inaccessible" was too strong, as several of these conditions are theoretically addressable by other wireless platforms. In the revised manuscript, we have replaced this phrasing

with: " Paired with a zero-shot behavior analyzer, Airscope enables cortical imaging across rich behavioral conditions, including room-scale enriched arenas, swimming assays, and multi-organ implantation, with concurrent recording from multiple animals."

Below we address each specific example:

4a — Room-scale enriched arenas

It is not clear why this experimental context would be inaccessible to existing wireless miniature microscope platforms (e.g., wire-free miniscopes). The authors should clarify what specific limitation of prior systems is overcome here.

Response: We thank the reviewer for this question. We acknowledge that, in principle, any wireless microscope system could be deployed in a room-scale arena. However, two practical considerations make Airscope particularly well-suited for this context:

1. Transmission range. Airscope's Wi-Fi-based transmission protocol maintains reliable connectivity across the full extent of the test arenas used in this study ($>5\text{ m}^2$, **Fig. R9**), supporting reliable operation in room-scale enriched environments.

2. Device weight and battery burden. Existing wireless miniscope platforms (e.g., battery-equipped Miniscope v3/v4) carry a substantially heavier combined device-plus-battery payload compared to Airscope, as listed in **Supplementary Table 2**. The reduced weight of Airscope minimizes motor impedance, thereby facilitating naturalistic exploration by the subject in topologically complex environments.

Accordingly, we have revised the manuscript to clarify that while room-scale imaging is not fundamentally exclusive to Airscope, specifically by removing the "inaccessible" characterization. We emphasize the distinct practical advantages that render such large-scale deployment reliably achievable with our system.

4b — Swimming assays

The manuscript demonstrates imaging during swimming by encasing the device to make it water resistant. It is unclear whether this capability is uniquely enabled by AirScope or whether similar encapsulation strategies could be applied to other miniature microscopes. Clarifying what aspects of the AirScope design make it particularly suitable for this application would strengthen this claim.

Response: We thank the reviewer for this insightful comment. We acknowledge that, in principle, waterproofing encapsulation methodologies could be adapted for other miniature microscope systems. However, the intrinsic design of Airscope confers two distinct practical advantages specifically beneficial for aquatic applications:

1. Low device mass. Airscope's low mass minimizes disruption to the animal's natural swimming kinematics. As detailed in Supplementary Table 2, all previously published wireless miniaturized microscopes have substantially greater payload than Airscope. The reduced mass lowers biomechanical burden during swimming and is the primary factor enabling stable head-mounting on a freely swimming mouse.

2. Practical implementation. We are not aware of prior demonstrations of a waterproof, head-mounted system used for calcium imaging during untethered swimming. While the waterproofing encapsulation technique itself is transferable to other devices, the combination of Airscope's low

mass, compact form factor, and fully wireless operation is what makes this application practical in vivo.

Accordingly, we have revised the manuscript to remove claims of absolute exclusivity and to instead emphasize the practical viability of Airscope for aquatic experimental paradigms.

4c — Multi-organ implantation

A total device mass of ~1 g is plausibly enabling for simultaneous implantation of multiple devices on a single animal. This claim appears reasonable and is supported conceptually, although additional discussion of practical constraints (e.g., spacing, mounting) would be helpful.

Response: We thank the reviewer for this constructive suggestion. To clarify, based on the physiological constraints outlined in the Mini2p⁵ study, the concurrent mounting of any existing wireless miniaturized microscope, with the sole exception of Airscope, onto both the cranium and spinal cord of a freely moving animal exceeds established weight-bearing tolerances (5g, **Supplementary Table 2**). Exceeding these payload thresholds inevitably induces abnormal motor behavior and compromises naturalistic activity. Furthermore, beyond these physical limitations, concurrent dual-site mounting and recording has not been demonstrated in any published literature utilizing other wireless miniaturized microscope platforms.

Accordingly, we have revised the manuscript to remove claims of absolute exclusivity, opting instead to highlight the distinct practical viability of Airscope for this demanding application.

4d — Multi-animal concurrent recordings

The manuscript suggests that AirScope uniquely enables simultaneous recordings from multiple animals. It is not clear why existing miniature microscope platforms could not, in principle, be extended to multi-animal operation. The authors should clarify whether this is a fundamental architectural advantage or an implementation choice.

Response: We thank the reviewer for this question. We wish to clarify that we do not claim fundamental exclusivity for multi-animal recordings. We acknowledge that, in principle, other wireless miniature microscope platforms could be extended to multi-animal operation.

The practical contribution of Airscope in this context is threefold: (a) the low device mass (~1 g) minimizes behavioral perturbation when multiple animals must wear devices simultaneously in socially interactive settings; (b) the Wi-Fi-based wireless architecture supports multiple concurrent devices without requiring dedicated receivers per animal (tested with 5 simultaneous units, see Non-Blocking Comment 5d below); and (c) the accompanying SAM2MICE behavioral analysis pipeline enables automated multi-animal tracking and individual identification without manual annotation.

The revised text explicitly acknowledges that multi-animal operation is not unique to Airscope but frames the demonstration as practical realization of simultaneous wireless cortical imaging during natural multi-animal social interactions, an experiment that is widely desired. We believe the practical enabling of this experimental paradigm constitutes a meaningful contribution to the field.

Non-Blocking Comment 5 — Claims of Innovation

5a — Aspheric optical design shortening optical conjugate by >50%

Aspheric optical design shortening the optical conjugate by >50% relative to prior head-mounted designs.

This claim is partially supported by Supplementary Fig. S2 and Supplementary Table 2, which illustrate how the custom aspheric elements contribute to a reduced conjugate length. This appears to be a genuine and interesting optical innovation.

Response: We sincerely thank the reviewer for recognizing this optical innovation. As described in our response to Non-Blocking Comment 3 and **Fig. R11**, we have expanded the quantitative comparison of optical track lengths as well as optical conjugates in the revised **Supplementary Fig. 1** and a new **Supplementary Note 1** to comprehensively describe the optical design procedure, providing explicit measurements and a systematic design comparison with prior systems.

5b — Space-bandwidth product comparison

Custom lens module preserving ~4 μm resolution across a 6 mm field of view and hemispheric working area, with increased space-bandwidth product.

The manuscript claims a doubling of space-bandwidth product relative to the highest-throughput wired miniscopes and an ~30 $\times$ improvement over prior wireless miniscopes. Including explicit space-bandwidth product values in Supplementary Table 2 would make this comparison clearer and more quantitative.

Response: We thank the reviewer for this suggestion. We have added explicit space-bandwidth product (SBP) values for all systems listed in the revised **Supplementary Table 1**. The table now includes: Effective FOV, measured resolution (center and edge), calculated SBP, frame rate, throughput (pixel readout rate per second), and spatiotemporal bandwidth product for each system. This enables readers to independently verify the claimed improvement factors.

5c — Integrated illumination, filtering, and mechanical stack

Integrated illumination, filtering, and mechanical stack enabling ~1 g total mass.

The integration of illumination, filtering, and mechanical components appears to be a key contributor to the low total mass. The illumination geometry itself seems innovative, but this aspect is not emphasized in the current description and could be highlighted more clearly.

Response: We sincerely thank the reviewer for recognizing this design element, which is indeed central to achieving the ~1 g total mass. In the revised manuscript, we have moved the description of the integrated illumination geometry into the first section of the Results, where it is now presented as a core design innovation alongside the optical system. The revised text, method section, and **Supplementary Fig. 2** clearly describe how the on-board LED illumination module, custom filtering, and mechanical stack were co-optimized to eliminate external fiber coupling — a key distinction from tethered systems that require fiber-delivered excitation light.

5d — Wireless architecture supporting high-fidelity data logging

The manuscript states that the wireless system supports high-fidelity data logging, but the specific performance metrics underpinning this claim (e.g., sustained bandwidth, packet loss, latency, buffering behavior) are not clearly documented. Given that total available wireless bandwidth must necessarily be shared across concurrently operating devices, it would be helpful to clarify how system performance scales as the number of active units increases. In particular, reporting whether bandwidth is dynamically allocated or statically divided, how packet collisions or dropouts are handled, and whether effective frame rate or data integrity are impacted during multi-animal experiments would strengthen this claim and aid interpretation of the reported results.

Response: We thank the reviewer for this thorough set of questions, which prompted us to document the wireless subsystem performance comprehensively. We have added a new Supplementary Figure and Methods subsection reporting:

1. Single-device performance metrics:

- Sustained bandwidth: 1238.74 KB/s (sufficient for 12.5 Hz UXGA MJPEG streaming)
- Packet loss rate: < 0.5% over 10 minutes recording
- End-to-end latency: 24.79 ms (MCU capture to host reception)
- Buffering: onboard 2 frames circular buffer handles transient Wi-Fi congestion

2. Multi-device scaling (1–5 simultaneous units):

We have performed systematic scaling tests with 5 concurrently recording Airscope devices on a single wireless access point. The results, presented in the revised **Supplementary Fig. 4**, demonstrate:

- **Bandwidth allocation and scaling:** Wireless bandwidth is dynamically allocated by the access point using standard Wi-Fi QoS mechanisms, rather than being statically divided among devices. Across 5 simultaneous devices, the system maintained a sustained bandwidth of 1196.73 ± 97.44 KB/s per device (**Fig. R12d**), sufficient to support 12.5 Hz UXGA MJPEG streaming. The effective frame rate remained stable at 12.41 ± 0.13 FPS (**Fig. R12a**), with no measurable frame-rate degradation.
- **Packet collisions/dropouts:** The Wi-Fi protocol handles collisions via standard CSMA/CA arbitration. Across the multi-device scaling tests, the observed drop rate remained low at $0.21 \pm 0.20\%$ (**Fig. R12b**), indicating that packet collisions or transient congestion did not measurably compromise data transmission.
- **Latency and buffering:** The measured latency is 26.49 ± 8.52 ms (**Fig. R12c**), remaining below the inter-frame interval required for 12.5 Hz acquisition. In addition, transient data-transfer fluctuations are absorbed by the onboard circular buffer, preventing short-term delays from causing substantial frame loss.
- **Data integrity:** Frame-level CRC verification confirmed that data integrity is maintained across multi-device configurations.

These results confirm that the wireless architecture supports reliable multi-animal operation without compromising data quality.

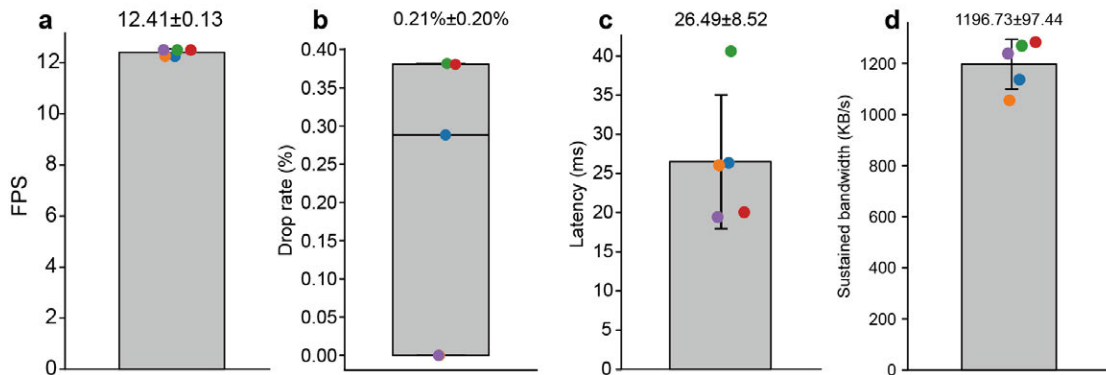


Fig. R12. Airscope wireless performance verification. **a.** FPS. **b.** Dropout rate. **c.** End-to-end latency. **d.** Sustained bandwidth. Values are presented as mean $\pm$ SD from five Airscopes, with all devices operating simultaneously on the same Wi-Fi access point.

We have added these verifications into **Supplementary Fig. 4** in the revised manuscripts.

Non-Blocking Comment 6 — Neuroscience claims (high-level summary)

I consider the neuroscience results at a high level as demonstrations of the types of experiments enabled by the AirScope platform. Accordingly, a detailed assessment of the experimental design, statistical analyses, and biological interpretations underlying these claims is outside my primary area of expertise, and I do not evaluate whether the stated biological conclusions are fully supported. Within this context, the manuscript presents several neuroscience applications enabled by the AirScope platform and makes the following biological claims:

Reward discovery during maze exploration The authors report the identification of a neural state associated with reward discovery during complex maze exploration, which they describe as previously uncharacterized (Fig. 5). **Spinal–cortical interactions during fear conditioning** Using simultaneous imaging of the spinal cord and cortex, the authors report interactions between spinal and cortical activity during fear conditioning and interpret these results as evidence of causal pathways linking these regions (Fig. 6). **Cortical representations of social hierarchy** The manuscript reports that allowing animals to freely interact and compete for resources reveals cortical representations of social hierarchy that are widely distributed across the brain within individual animals (Fig. 4). **Multiplexed representations of reward expectancy and navigation** In complex foraging mazes, the authors report distributed cortical activity patterns that multiplex reward expectancy with spatial navigation variables (Fig. 5).

Response: We thank the reviewer for this transparent and clear summary of the biological claims. We appreciate the reviewer's candor regarding expertise boundaries and fully agree that these claims require evaluation by neuroscience-expert reviewers.

Based on comprehensive feedback from Reviewers 2 and 3 (who evaluated the biological claims in depth), we have made the following revisions:

1. Language precision. All biological claims have been systematically revised to use observational rather than causal language (e.g., "identify" $\rightarrow$ "observe," "causal pathway" $\rightarrow$ "Granger-causal influence," "unveil cortical basis" $\rightarrow$ "demonstrate cortical imaging during").

Confounding variables (motor actions, sensory experience, low-level visual features) are now explicitly acknowledged for each experiment.

2. Focused flagship experiment. The revised manuscript focuses on one flagship neuroscience demonstration — the social competition experiment in freely behaving animals. This section has been strengthened with population-level analyses showing that dorsal cortical activity carries information about competitive behavior. We further observed coordinated cortical dynamics between interacting animals, consistent with previously reported⁶. We have also added controls for potential motor confounds, supporting the interpretation that these signals are not simply explained by movement or behavioral state. Secondary neuroscience demonstrations are presented more concisely with appropriately tempered claims.

3. Enhanced transparency. Raw data visualizations, step-by-step analysis pipeline outputs, and two-photon ground-truth comparisons are now provided for the key experiments, enabling readers to independently assess data quality at each stage.

Minor Comments (Figure-Level)

Minor Comment 1 — Figure 1 revisions

Panel a: This panel does not appear necessary for a scientific readership and may be more appropriate for a lay audience. Consider removing it or moving it to the Supplementary Information.

Response: We thank the reviewer for this suggestion. We have considered removing this panel but ultimately decided to retain it, as we believe it serves an important narrative function. Panel a establishes the core motivation — that wireless miniature microscopy opens access to naturalistic behaviors inaccessible to tethered systems — for readers across disciplines who may not be familiar with the constraints of current head-mounted imaging. We have, however, refined the panel to improve its visual clarity and scientific tone, removing elements that may have appeared overly promotional.

Specifically, we removed the large directional arrow, as it introduced a promotional, advertisement-like visual flow inconsistent with the conventions of scientific figures. In the revised panel, the two conditions — Prior approaches (head fixation and tethered observation) and Airscope (wire-free, fully unconstrained observation) — are now presented as a side-by-side comparison. This layout more clearly frames the figure as a schematic contrast between existing methodological constraints and the capabilities enabled by the present system, rather than implying a progressive or sales-oriented narrative. We believe this revision better serves a scientific readership while retaining the conceptual motivation for the work.

Panel c: This panel presents the optical design; however, the corresponding subpanel is not clearly described in the figure caption. Additional description would improve clarity.

Response: We thank the reviewer. We have expanded the figure caption for panel c to clearly describe each subpanel element, including the ray trace diagram, lens surface profiles, and the design specifications displayed.

Panel e (Filtering strategy): The caption states that modeling and optimization minimize spectral leakage while maintaining fluorescence collection efficiency. While this statement is plausible, it is not clear how this claim is supported by the image shown. Clarifying the connection between the modeling and the presented panel would strengthen this section.

Response: We thank the reviewer for this observation. The full optimization model and spectral leakage analysis are described in the Methods section and a dedicated **Supplementary Fig. 3**. Panel e is intended as a simplified visual summary. We have refined the panel to better illustrate the logical connection between the spectral modeling and the resulting filter design, including a brief annotation linking to the Supplementary Fig. 3 where the complete analysis is presented.

Panel f: This panel may be better suited for the Supplementary Information, as it does not appear essential to the main narrative.

Response: We agree with the reviewer. Panel f has been moved to the **Supplementary Fig. 3** and the remaining panels have been reorganized accordingly to improve the main figure's coherence.

Panel g (Wireless system architecture): The caption describes a schematic of the wireless mode and an example of brain-wide activity data. As written, the panel does not clearly communicate what specific aspects of the wireless architecture or performance are being demonstrated.

Response: We thank the reviewer for this helpful comment. We revised panel g (now Panel f) to more clearly illustrate the wireless imaging workflow and explicitly highlight the role of Wi-Fi-based transmission in enabling real-time remote image preview and data streaming. The updated schematic now depicts the complete pipeline, including Wi-Fi transmission from the Airscope to the receiver, real-time image preview, imaging frame acquisition, and extraction of $\Delta F/F$ activity traces. We also updated the caption to clarify the data path from the Airscope to the receiver and the corresponding activity readout.

The caption now reads like: “*f, Wireless imaging workflow. Airscope transmits imaging data to the receiver through Wi-Fi for real-time remote image preview and wireless data streaming. After preview-based confirmation, imaging frames are acquired and processed to extract neuronal $\Delta F/F$ activity traces over time.*”

“Snap-and-go” concept: This is a compelling and innovative idea, but the current figure presentation does not fully convey its significance. Consider revising the figure or caption to better illustrate this functionality.

Response: We sincerely thank the reviewer for recognizing the significance of the snap-and-go design. We have revised the figure and the corresponding caption to better illustrate this functionality as a step-by-step workflow, including rapid battery attachment, wireless recording, quick battery exchange, and immediate initiation of a new session. We also added timing annotations to highlight the speed and simplicity of the procedure, showing that the battery can be attached in less than 1 s and exchanged in less than 2 s with minimal disruption to the animal.

Panel grouping: Panels *c, d, and i* could be grouped more coherently to present a unified optical and system design narrative. Similarly, panels *b, j, and k* could be grouped to improve flow and readability.

Response: We thank the reviewer for this constructive organizational suggestion. We agree that segregating the system schematics from the performance metrics presents a more cohesive narrative. Accordingly, we have comprehensively restructured Figure 1 into four distinct thematic rows to optimize logical flow and readability:

- Row 1 (Panels a–b): Concept and Throughput. Introduces the Airscope concept and highlights its superior neuronal recording throughput compared to existing wireless systems.
- Row 2 (Panels c–e): Optical Design. Details the optical architecture (c), illumination optimization (d), filtering optimization (e). Panel f has been removed as requested.
- Row 3 (Panels f–h): Wireless Architecture and Assembly. Outlines the wireless data transmission scheme (f), the rapid-exchange battery protocol (g), and the overall system assembly (h).
- Row 4 (Panels i–o): System Performance Benchmarks. Compares the space-bandwidth product (SBP), weight, height, and field of view (FOV) against alternative systems (i–k), and demonstrates motion agility (l–o).

We believe this revised layout substantially enhances the figure's clarity and pedagogical value.

Figure resolution: Given the high number of subpanels, increasing image resolution would improve readability. Several subpanels appear pixelated when zoomed, which makes detailed inspection difficult.

Response: We thank the reviewer. In this revision, we have increased the figure resolution to 300 ppi across all panels, ensuring that all subpanels remain crisp under magnification.

Panels m–p: These panels focus on device performance rather than device construction. They provide useful characterization, but their relationship to the rest of the figure could be clarified.

Response: We thank the reviewer. Panels m–p are intended to position Airscope within the landscape of existing miniaturized microscopes through quantitative performance benchmarking. We have revised the figure configuration to enhance this transition from device construction (upper panels) to performance comparison (lower panels).

Use of arrows: The use of parabolic arrows in the figure does not add clarity and may come across as promotional. Simplifying or removing these elements would improve the scientific presentation.

Response: We appreciate the reviewer's comment. We would like to clarify that the arrows were not intended as promotional elements, but rather to guide the reader's attention to the overall nonlinear trend of the data. In both panels j–k, the distribution of points does not follow a simple linear trajectory. Therefore, a curved arrow helps indicate the direction of this trend more clearly than a straight line or no guide at all.

To address the concern, we have simplified the visual style of the arrows by making them less prominent, while retaining the curved trend indication to preserve the scientific interpretation of the data.

Minor Comment 2 — Figure 2 revisions

Panel a0: The comparison between the AirScope system and a microscope is not particularly informative, as the microscope's substantially larger size inherently enables higher performance. This comparison may be misleading or unnecessary.

Response: We thank the reviewer for raising this concern, and we agree that a head-to-head performance comparison with a benchtop microscope would be misleading, since the microscope's substantially larger optical volume inherently affords higher resolution. This was never our intent.

To avoid this implication, we have revised panel a and its caption to frame the comparison as a qualitative reference rather than a performance benchmark. Specifically, panel a now shows Airscope ($13 \times 13 \times 8.2 \text{ mm}^3$) and a benchtop microscope ($50 \times 50 \times 30 \text{ cm}^3$) imaging the same preparations side-by-side: neuronal fibers in the spinal cord (panel b) and the dorsal cortex (panel c). The purpose is solely to demonstrate that, despite its compact form factor, Airscope image quality is sufficient for mesoscale neuronal imaging in single-cell resolution. We do not claim equivalent optical performance at reduced size, and the revised caption now states this intent explicitly.

Panel c: The Fourier transform power density inset is a useful and informative addition.

Response: We thank the reviewer for this positive assessment. We have retained the Fourier transform power density inset for the neuron (**Fig. 2c**) and microglia (**Supplementary Fig. 9a**) imaging results.

Reviewer #2

Remarks to the Author

The manuscript by Zhang and coworkers describes ‘Airscope’, a wireless mesoscope epifluorescence system featuring a 6 mm diameter field-of-view with lateral 4 – 5 micrometer optical resolution and 10 Hz frame rate, weighting 1 g. The 1.7 g supplied battery enables 15 minutes autonomy and can be rapidly swapped to continue recording. An alternative wired power line enables long uninterrupted recordings. The authors describe a fully automated analysis pipeline featuring SAM2MICE, a behaving-mice segmentation algorithm for automated videography analysis based on foundational model SAM2. Together these techniques are used in a wide range of behavioral paradigms including social interaction, swimming and large environment exploration. These impressive technical advances enable a robust and versatile device, which linked to an open-sourced release promises to have an impact in neuroscience of freely-moving animals if a few of the details of the microscopes capabilities and caveats are demonstrated fully.

We thank the reviewer for this clear and accurate summary of the Airscope platform, and for recognizing the engineering advances, the versatility across behavioral paradigms, and the value of the open-source release. We are particularly grateful for the reviewer's framing, that Airscope is poised to impact freely-moving neuroscience *once* the microscope's capabilities and caveats are demonstrated fully.

While I applaud the authors on their engineering feat and potential application of the microscope, at present many of the major claims of the paper are somewhat unjustified as they are presented and need addressing. Overall, the manuscript is very unfocussed and moves from the new technique to behavior to brain activity correlates and behavior. While I understand that the aim was to show the many different potential applications that the new approach could be applied to, the issue is that it is hard to understand if any of the biological claims are true and whether they would hold up to deeper scrutiny. This leads to the biological results being very descriptive and relying on ‘black-box’ analysis techniques to make any sense of the data but leaves the reader unable to really assess the quality of the raw data being gathered by the microscope and the subsequent step by step analysis steps.

Response: We thank the reviewer for this candid and constructive critique. We took the central concern, that the manuscript moved too quickly across techniques, behaviors, and biology, leaving the reader unable to assess either the raw data quality or the individual analysis steps, as the organising principle for this revision. The result is a substantially more focused manuscript anchored to a single neuroscience question that Airscope uniquely enables, supported by interpretable analyses and a fully public dataset and pipeline. We summarise the major changes here and elaborate on each in the point-by-point responses below.

1. Sharpened biological scope. The multi-animal experiments have been refocused and substantially expanded around a single concrete question: *is competitive rank reflected in distributed cortical population activity beyond prefrontal cortex, and does the inter-brain neural coupling reported in medial prefrontal cortex⁴ extend to the wider dorsal cortex?* Both formulations test whether prefrontal findings reflect a broader principle of cortical organization, and neither can be addressed with existing tools: tethered miniscopes preclude two unrestrained animals from freely encountering each other in tubes, single-port feeders, or running wheels, and prior wireless devices are too bulky for paired imaging. Simultaneous large-scale dorsal-cortical

recording from two freely competing mice — what Airscope makes possible — is precisely the experiment that lets these questions be asked, and it now anchors the biological narrative of the paper. Secondary biological demonstrations have been consolidated into the Supplementary Information.

2. Independently verifiable raw-data quality. To allow direct external evaluation of the data Airscope produces, we built a hybrid system in which the same cortical region in the same animal is recorded simultaneously by Airscope and by a benchtop two-photon microscope serving as ground truth. Across 36 paired recordings in three mice (1,196 neurons), Airscope achieved an F1 score of 0.78 ± 0.05 , a recall of 0.81 ± 0.08 , and a temporal-trace correlation of 0.82 ± 0.14 against two-photon ground truth (**Major Comments 3 and 4**). In parallel, the optical characterization — full-FOV resolution map, USAF and fluorescent-grid FOV calibration, MTF, transmission bandwidth, and a head-to-head specification table against the Kiloscope — has been substantially expanded (**Major Comment 1**; see also Reviewer 1, Blocking Comments 1–3).

3. Fully open data and pipeline. Following **Major Comment 7**, the `Airscope_toolbox` repository is now public and includes raw recordings, the calcium-imaging processing pipeline, SAM2Mice behavior segmentation, the decoding code, documented inclusion criteria, and end-to-end Jupyter-notebook tutorials. Any reader can reproduce the segmentation, trace extraction, and downstream analyses step by step from the raw frames.

4. Interpretable analyses replace black-box readouts. Where the original manuscript leaned on opaque decoders, the revised analyses are hypothesis-driven and resolved at the single-cell level, including PCA-based alignment of population geometry across animals, cell-resolved inter-brain correlation maps, and SVM cross-condition generalization tests. Motor and sensory confounds are now explicitly addressed for each behavioral paradigm (**Minor Comment 4**; see also Reviewer 3, **Major Comment 3**).

5. Tightened language across the manuscript. Causal and mechanistic phrasing has been replaced with observational language throughout (e.g., "*encode*" → "*correlate with*").

In addition, the revision period produced three further capabilities that touch on the reviewer's other comments: a paired tethered-vs-wireless comparison demonstrating behavioral and neural activities differences between wireless recordings and tethered recordings (**Major Comment 5**); a wireless 30 Hz high-speed variant (**Minor Comment 6**); and two optogenetic-capable Airscope variants — a side-illumination version with separately controllable blue (imaging) and red (inhibition) LEDs, and an Airscope-plus-microLED implant configuration which we used to demonstrate wireless seizure suppression (**Minor Comment 4** and **Major Comment 5**).

We hope these changes, together with the detailed point-by-point responses below, address the reviewer's structural concern while making the underlying data and analyses fully open for independent step-by-step assessment.

Major comments:

Major Comment 1 — Novelty relative to Kiloscope (injection-molded aspheric lenses)

1. If optics only are considered, previous work by Sherrer et al. already published an optical system based on injection-molded aspheric lenses. This previous system is well acknowledged by authors and a quantitative comparison is available, presenting an improvement by the new design (factor 2 of space bandwidth metric), which overall is an incremental advance and should be presented as

a solid step but not a major innovation and certainly was not the first to apply this approach to lightweight 1P microscopes.

Response: We thank the reviewer for raising this important point and for the careful comparison with the Kiloscope system by Scherrer et al⁷. We respectfully disagree that Airscope represents merely an incremental optical advance, and we would like to clarify our contribution from three perspectives:

1. Optical performance differences. The Airscope optical architecture is not only 2.84× shorter in total track length (TTL) than the Kiloscope, but it also yields a more than twofold increase in spatial bandwidth. Furthermore, the Airscope design exhibits distinct advantages across other critical system parameters, including extended working distance and streamlined optomechanical integration. A comprehensive comparison of these optical specifications is provided below.

Table R3. Comparison of design parameters of Airscope and Kiloscope. Airscope achieves a more compact and integrated design while providing a longer working distance, improved MTF, and higher optical bandwidth.

Metric	Airscope	Kiloscope	Advantages
Track length	3.36 mm	9.55 mm	~65% shorter
Conjugate length	7.52 mm	10.89 mm	~31% shorter
Working distance	4.16 mm	1.35 mm	~3.1× longer
Glass volume	0.016 cm ³	0.063 cm ³	~75% less
Integration/assembly	Already integrated	Need manual integration/potential to introducing additional aberrations	Integrated design reduces assembly complexity and alignment-induced aberrations
MTF @ 250 lp/mm 0.5 FOV	0.33	0.16	~2.06× higher
Bandwidth	4.68 × 10 ⁶	1.92 × 10 ⁶	~2.44× higher

2. Design scalability. While both systems employ aspherical lenses, their underlying design philosophies differ fundamentally. The Kiloscope repurposes off-the-shelf cellphone lens modules, which constrains the achievable field of view to the native sensor–lens pairing and inherently limits architectural flexibility. In contrast, the Airscope utilizes custom-designed aspherical optics optimized specifically for mesoscale cortical imaging. Specifically, comprehensive design details are now provided in new **Supplementary Note 1**. As illustrated in **Fig. R13b**, we optimized the Airscope framework using a custom curriculum-learning optical design agent, termed LensCopilot. LensCopilot first generates a structured multi-stage optimization curriculum in JSON format that serves as a fixed blueprint for the design process. The curriculum defines a sequence of stages with progressively stricter targets, beginning from conservative settings, such as a narrow field of view and a large F-number, and gradually approaching the final specifications. This curriculum-learning

strategy improves optimization stability and reduces the risk of early convergence to poor local minima. A key feature of the curriculum is the scheduled release of degrees of freedom (DoFs), whereby variables such as surface radii, thicknesses, and aspheric coefficients are unlocked sequentially at predefined stages. Optimization is initially restricted to low-order aspheric terms, particularly A4, to establish the basic lens geometry and first-order optical performance, while higher-order terms, such as A6 and A8, are introduced later to correct more complex aberrations. The curriculum further enforces dynamic manufacturability constraints, including minimum edge thickness and TTL, and uses reusable merit-function templates to ensure deterministic and consistent evaluation throughout the optimization trajectory. A deterministic executor then parses the curriculum and performs the optimization in Zemax OpticStudio using a predefined local–Hammer–local (LHL) loop at each stage: a bounded damped-least-squares local optimization, followed by a Hammer-based global search under a fixed iteration budget, and concluding with a constrained local refinement. The final optimized optical design of the Airscope is detailed in **Extended Data Fig. 1**.

This distinction is critical: the Kiloscope architecture is fixed, and cannot readily support a 6 mm FOV or beyond, whereas the Airscope design framework is inherently scalable. To see this, we have now demonstrated by producing additional configurations (see figure of **Supplementary Note. 1**). In these additional setups, Airscope optimization-styles produces two variants: 1) Centimeter-scale FOV system, keep single-cell resolution (5 μm); and 2) Smaller FOV system with high resolution (1 μm). We believe providing a *generalizable design methodology* for the community constitutes a significant contribution beyond any single device.

3. Special optical design-enabled on-board illumination. The optical architecture of the Airscope enables systematic system-level optimization, notably the integration of on-board illumination that obviates the need for the external optical fiber required by the Kiloscope. Specifically, the Kiloscope's reliance on fiber-coupled illumination and tethered data readout inherently precludes fully wireless operation (**Fig. R13f**). In contrast, the Airscope features a substantially expanded working distance (the spatial clearance between the sample and the first optical element). This extended clearance provides the necessary physical volume to seamlessly integrate LEDs. By leveraging this structural allowance in conjunction with a customized illumination optimization algorithm (**Supplementary Fig. 2**), the Airscope achieves robust, on-board illumination across a comprehensive 6-mm field of view (FOV).

4. System-level integration. The core innovation of Airscope is not optics in isolation but the co-optimization of optics, wireless data transmission, on-board illumination, and miniaturized form factor into a fully untethered system weighing only 1 g. The engineering challenge of maintaining high spatial bandwidth while simultaneously achieving wireless capability, on-board LED illumination, and sub-gram optical payload has not been previously addressed. We have revised the manuscript to more clearly articulate this system-level contribution and to present the optical advance as one component of a broader integrated innovation.

We have added the above optimization details as new **Supplementary Note 1**.

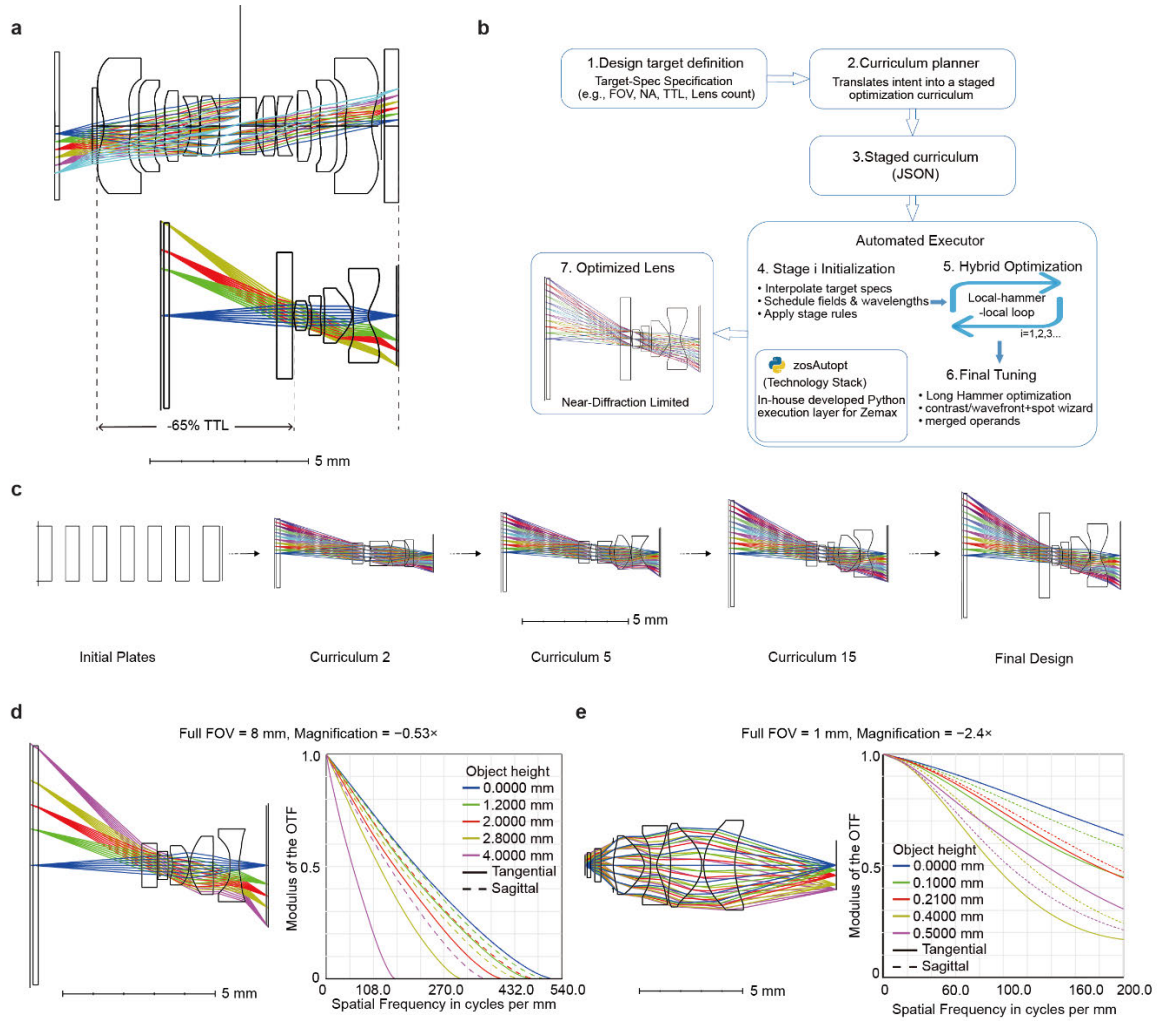


Fig. R13. Comparative Analysis of Kiloscope and Airscope Architectures. **a**, Ray-tracing schematics (Zemax) comparing the Kiloscope (top) and Airscope (bottom) optical configurations.

Airscope achieves a $2.84\times$ shorter total track length (TTL) and much increased field of view (FOV) relative to the Kiloscope. **b**, Schematic representation of the Airscope optimization workflow utilizing a proposed curriculum learning framework (Supplementary Note 1). This proprietary design engine enables arbitrary geometric configurations. **c**, Evolution of the Airscope design, transitioning from a planar parallel-plate initial state to the final optimized form (left to right). **d**, Example of an additional four-element lens generated using the same automated framework for a wide-field configuration (8 mm FOV, $-0.53\times$ magnification). **e**, Example of another four-element lens generated using the same framework for a higher-resolution, higher-magnification configuration (1 mm FOV, $-2.4\times$ magnification). Together, panels **d–e** illustrate the scalability of the design methodology beyond a single lens instance. **f**, Cross-sectional comparison of the prior Kiloscope system⁷ (left) and Airscope (right). The prior system uses fiber-coupled illumination and a prism-based opto-mechanical arrangement, whereas Airscope adopts integrated on-board LED illumination and a compact co-designed architecture combining optics, filters, mechanics, electronics, and packaging. The Kiloscope cross-section is adapted from Scherrer *et al.* (2023)⁷. Scalebar: 5 mm (**a**, **c**, **d**, **e** and **f**).

Major Comment 2 — Time-space bandwidth product comparison

2. The space-bandwidth metric is well suited to evaluate optics, but does not include temporal resolution. In the same previous work, the mesoscope operated at 30 Hz frame rate, whereas Airscope performs at 10 Hz. A more revealing of overall performance of the whole system (not just optics) would be a time-space bandwidth metric, which would be relatively close between the kiloscope from Sherrer et al and Airscope.

Response: We thank the reviewer for this constructive suggestion. We address it from three complementary angles: biological adequacy, quantitative comparison, and achievable frame rate.

1. Biological adequacy of the 10 Hz frame rate

For in vivo calcium imaging with GCaMP6f, the indicator kinetics (decay constant over 230 ms)⁸ are sufficiently captured at a 10 Hz sampling rate. To empirically validate this, we utilized publicly available two-photon imaging datasets from the Allen Brain Observatory (ABO; <https://observatory.brain-map.org/visualcoding>), which were originally acquired at 30 Hz within comparable cortical regions, rather than data acquired in this study. The temporal data were subsequently downsampled to 10 Hz. Qualitative assessment revealed nearly identical fluorescence profiles between the two sampling rates. Quantitative analysis confirmed that no significant calcium transient information was lost at the lower frame rate (**Fig. R14**). This was evidenced by high temporal fidelity at the trace level (Pearson correlation coefficient, mean $\pm$ SD: 0.9856 ± 0.0315) and consistency in estimated neural activity. Specifically, spike deconvolution performed via the OASIS algorithm showed no significant differences in inferred firing rates (Wilcoxon rank-sum test, $p = 0.205$).

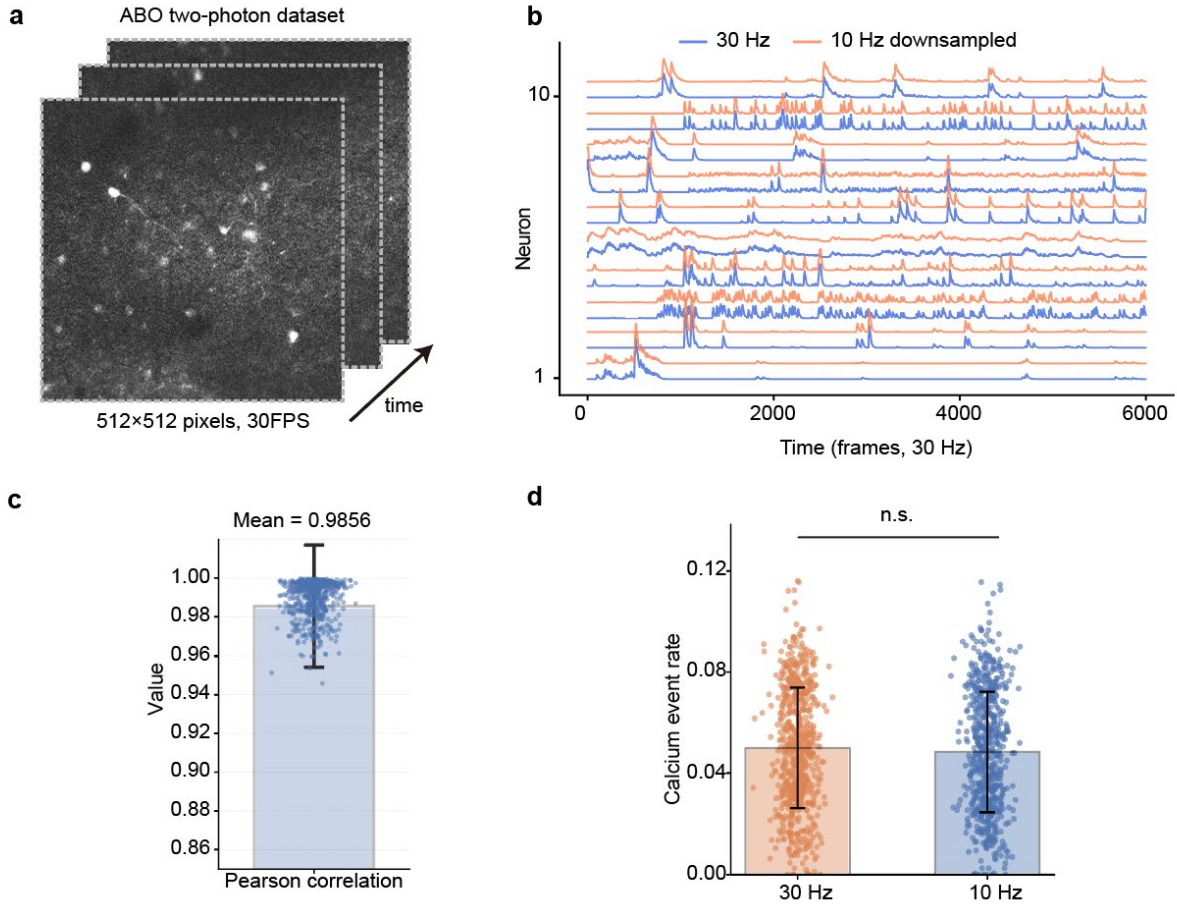


Fig. R14. Validation of 10 Hz Sampling Rate in Rasgrf-Ai148D Transgenic Models. **a**, Representative two-photon calcium imaging datasets acquired from Rasgrf-Ai148D mice at a native frame rate of 30 Hz. **b**, Qualitative comparison of fluorescence traces between the original 30 Hz acquisition (blue) and the 10 Hz temporally subsampled data (orange). **c**, Pearson correlation coefficients calculated across 24 sessions, $n=664$ neurons, comparing native (30 Hz) and downsampled (10 Hz) temporal profiles (Mean $\pm$ SD: 0.9856 ± 0.0315 ; error bars represent SD). **d**, Comparative analysis of calcium event rates. Statistical significance was assessed using a two-sided Wilcoxon rank-sum test ($p = 0.205$, indicating no significant difference).

2. Quantitative comparison with Kiloscope

We address the reviewer's space-bandwidth metric concern from two perspectives:

(a) Spatial bandwidth product (SBP). The SBP is the definitive measure of a system's optical performance. Airscope achieves an SBP of 4.68×10^6 , which is **$\sim 2.44\times$ higher** than the Kiloscope's 1.92×10^6 . This reflects Airscope's superior optical design.

(b) Spatiotemporal pixel throughput. We note that the 10 Hz rate used above reflects the acquisition setting for the biological-adequacy validation; through subsequent firmware optimization, the Airscope now operates at a frame rate of 12.5 Hz, which we adopt for the throughput and bandwidth comparisons below. Following the reviewer's suggestion, we now

include a time–space bandwidth comparison in the revised manuscript. The throughput is calculated as the per-second pixel throughput:

Airscope: $1600 \times 1200 \text{ pixels} \times 12.5 \text{ Hz} = 2.4 \times 10^7 \text{ pixel} \cdot \text{Hz}$

Regarding the Kiloscope system, the authors did not explicitly report frame rate specifications in their original publications. However, analysis of their Methods section reveals the utilization of the Miniscope V3 Data Acquisition (DAQ) hardware. According to the technical specifications provided in the Miniscope V3 DAQ repository (<https://github.com/Aharoni-Lab/Miniscope-DAQ-Cypress-firmware>), the spatiotemporal pixel throughput can be derived using the following relationship:

Kiloscope: $800 \times 600 \text{ pixels} \times 30 \text{ Hz} = 1.44 \times 10^7 \text{ pixel} \cdot \text{Hz}$

Thus, even when temporal resolution is factored in, Airscope maintains a **$\sim 1.67\times$ higher information throughput** (a 67% lead) compared to this tethered microscope, while simultaneously delivering $2.44\times$ the spatial bandwidth product.

(c) Spatiotemporal bandwidth product (STBP).

Considering the frame rate of 30 fps for the Kiloscope and 12.5 fps for the Airscope, the spatiotemporal bandwidth comparison is **Airscope: $5.85 \times 10^7 \text{ Hz}$, while Kiloscope $5.76 \times 10^7 \text{ Hz}$** . However, we would like to note that the above comparison is between a wireless system (Airscope) and a wired system (Kiloscope), which makes Airscope's advantage all the more significant. Compared with other wireless systems (Miniscope Wirefree, Wireless miniScope and Finchscope), Airscope leads by $9.15\times$ STBP.

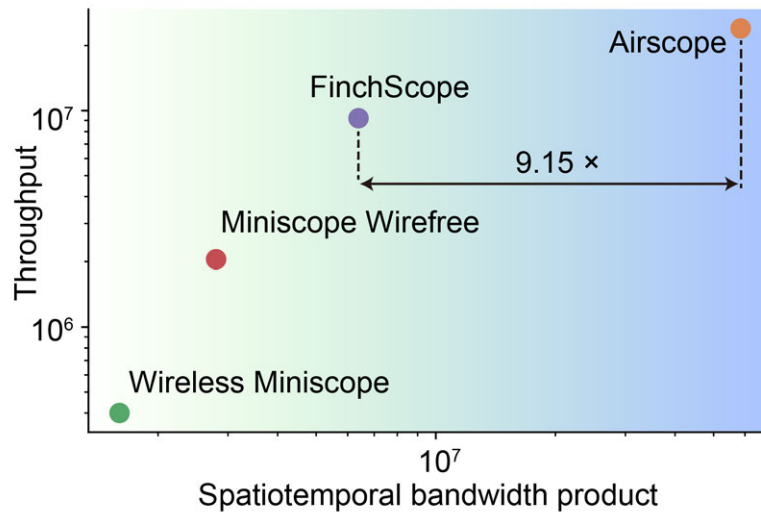


Fig. R15. Benchmarking of system throughput and spatiotemporal bandwidth product against state-of-the-art wireless miniaturized microscopes.

In the revised manuscript, we have supplied both spatial bandwidth product and spatiotemporal bandwidth product for comparison across different miniaturized microscopes, as required by the reviewer.

Year	Name of techniques	FOV (mm)	Resolution (μm)	Frame Rate (Hz)	Weight (g)	Height (mm)	Width (mm)	Volume (cm^3)	Throughput (pixels/s)	Spatial bandwidth product	Spatiotemporal bandwidth product
2017	FHIRM-TPM ⁹	0.13×0.13	0.64	40	2.15	20	16.25	3.25	2.6×10 ⁶	1.65×10 ⁵	6.60×10 ⁶
2024	TINIScope ¹⁰	0.45×0.45	3	40	0.43	8.24	13.8	0.52	5.18×10 ⁶	9×10 ⁴	3.60×10 ⁶
2018	CHEndoscope ¹¹	0.5×0.5	2	20	4.5	20	NA	NA	5.54×10 ⁷	2.5×10 ⁵	5.00×10 ⁶
2022	MINI2P ⁵	0.5×0.5	1.15	15	2.4	25	15.2	5.78	9.83×10 ⁵	7.56×10 ⁵	1.13×10 ⁷
2011	Integrated Miniscope ¹²	0.8×0.6	2.5	36	1.9	20	11	2.4	1.11×10 ⁷	3.07×10 ⁵	1.11×10 ⁷
2016	UCLA Miniscope ¹³	0.8×0.6	3	30	3	22	12	1.98	1.08×10 ⁷	2.13×10 ⁵	6.39×10 ⁶
2018	UCLA Miniscope v3 ¹⁴	0.70×0.45	3	30	3	22.5	13	1.98	1.08×10 ⁷	1.4×10 ⁵	4.20×10 ⁶
2020	Miniscope Wirefree ¹⁵	0.7×0.45	3	20	3.9	24	17.05	2	2.05×10 ⁶	1.4×10 ⁵	2.80×10 ⁶
2018	Doric μ Scope	0.73×0.73	2	45	3.7	16.6	13.9	2.52	1.79×10 ⁷	5.33×10 ⁵	2.40×10 ⁷
2014	InscopinVista ¹⁶	0.75×0.9	2.08	20	2.2	22	15	2.28	1.92×10 ⁷	6.24×10 ⁵	1.25×10 ⁷
2020	NINscope ¹⁷	0.79×0.50	3	30	1.6	18.2	11	2.2	1.44×10 ⁷	1.67×10 ⁵	5.01×10 ⁶
2017	FinchScope ¹⁸	0.8×0.6	3	30	1.8	16	8.5	1.16	9.22×10 ⁶	2.13×10 ⁵	6.39×10 ⁶
2019	Wireless Miniscope ¹⁹	0.5×0.5	2.5	10	3.9	30	18	8.64	4.00×10 ⁵	1.6×10 ⁵	1.6×10 ⁶
2021	UCLA Miniscope v4 ²⁰	Φ 1.0	3	30	2.6	22.3	16.9	5.88	1.08×10 ⁷	3.49×10 ⁵	1.05×10 ⁷
2023	Featherscope ⁷	1.0×1.0	4-7	30	1	15	16.5	1.24	1.44×10 ⁷	2.5×10 ⁵	7.50×10 ⁶
2023	MiniLFOV ²¹	3.6×2.7	3.5–6.2	11	13.9	38	40	12	5.54×10 ⁷	1.97×10 ⁶	2.17×10 ⁷
2025	MINIXL ²²	Φ 3.5	4.4	23	3.5	30	15	7.5	2.3×10 ⁷	1.99×10 ⁶	4.58×10 ⁷
2023	Kiloscope ⁷	4.8 × 3.6	6	30	1.4	15	17	2.04	1.44×10 ⁷	1.92×10 ⁶	5.76×10 ⁷
2025	Airscope	Φ 6	4-6	12.5	1	8.15	13	1.45	2.40×10 ⁷	4.68×10 ⁶	5.85×10 ⁷

Supplementary Table 1 | A comparison of open-source and commercial devices, highlighting key optical, digital and mechanical specifications. We sort miniaturized microscopes based on their field of views (FOVs). Throughput (pixels s⁻¹): total pixels transmitted per second (frame × width × height / frame interval). Spatial bandwidth: FOV (mm²) divided by square of half of lateral resolution (μm^2), reflecting the total resolvable area (Methods). Spatiotemporal bandwidth product is calculated by the product of spatial bandwidth and frame rate. NA: not available or not applicable. The computation of space bandwidth product of miniaturized microscope has taken into account the differential lateral resolutions throughout the entire field of field²³.

Major Comment 3 — Ground truth for signal fidelity in 3D

3. The single-cell resolution claimed in the manuscript is unjustified. Cellular resolution is in 3D not just 2D, and while the lateral resolution is enough for lateral resolution this is not the case in the axial plane. Given the tight packing of neuronal somas in the cortex in both lateral and axial planes, how can the activity from neurons stacked in the axial plane be separated? This is especially

important when considering the biological claims later in the manuscript where single-cell-resolution are required.

Response: We thank the reviewer for this important question. We understand the reviewer's concern that the 3D distributed neurons will eventually contribute to the mix signal of Airscope detection, and hamper the single-cell detection.

1. We acknowledge that the 90-104 μm depth of field allows us to capture more neurons (**Supplementary Fig. 8**), but it also introduces the potential issue of crosstalk from neurons in the axial plane. However, our algorithm ensures reliable demixing as long as neurons are not fully overlapped laterally. This algorithm incorporates DeepWonder²⁴ with a specialized demixing feature, designed to delineate spatially co-localized but temporally disparate entities, such as neuron segments that manifest in distinct temporal frames. To quantify and verify the cell demixing ability of DeepWonder, Zhang et al simulated three laterally overlapped neurons, each with distinct temporal activities are simulated. The neuronal distance was calculated by the distance between the center of mass of two neurons. Zhang et al then evaluated the correlation scores of neurons as a function of neuronal distance. They found that the correlation scores remained high even when two neurons were as close as 0.3 neuron diameters in **Fig. R16**.

[Figure Redacted]

Fig. R16. Neuron demixing in DeepWonder (Supplementary Fig. 6 from DeepWonder article). **a**, Segmentation results of DeepWonder, compared with ground truth. Three Gaussian-shaped neurons with the same diameters are simulated. Top two rows: The footprints of three neurons with a center distance of 0.3 neuron diameter and the calcium signal traces extracted by DeepWonder. Bottom two rows: The results of three neurons with a center distance of 0.7 neuron diameter. All the footprints and calcium transients match well with the ground truth. The neuronal distance is calculated by the distance between the center of mass of two neurons. The normalized distance (norm distance) is defined by the ratio of the neuronal distance over the neuron diameter. **b**, The correlation scores of calcium signals change with increasing neuronal distance. **c**, The Intersect over Union (IoU) of footprints in the presence of increasing neuron center distance. Statistical scores are shown in mean $\pm$ SD across $n = 3$ samples.

2. To further demonstrate the single-cell resolution and the signal fidelity of Airscope recording, we establish ground truth by building a hybrid system combining Airscope with a two-photon (2p) microscope, which is the gold standard for neuronal recording. In this setup, we coupled Airscope to a 2p microscope and temporally interleaved Airscope detection under volumetric one-photon illumination with 2p imaging, enabling us to obtain 2p ground truth data. The details of the system setup are shown as follow, in **Fig. R17a**.

For the Airscope path, a 1:1 magnification optical path was introduced before the Airscope system, with a 470-nm centered LED (M470L5, Thorlabs) used for excitation, and Airscope served as the relay imaging module to form an equivalent image. For the two-photon arm, a 920-nm excitation and detection path were incorporated into the system.

For the two-photon excitation path, a femtofiber 920-nm laser system (Femtofiber ultra 920, TOPTICA) served as the two-photon excitation source (920 nm central wavelength, pulse width <100 fs, 80 MHz repetition rate). A 4f system was used to expand the laser beam to a resonant scanner (8315K/CRS8K, Cambridge Technology). The scanned beam went through a scan lens (SL50-2P2, Thorlabs) and a tube lens (TTL200MP, Thorlabs) and formed a tight focus through an objective lens (TL10X-2P, Thorlabs).

To separate the two-photon excitation beam with one-photon 470-nm centered excitation beam and fluorescent emission path, a long-pass dichroic mirror (DMLP650L, Thorlabs) was introduced. Then, we put the second dichroic after the two-photon long-pass dichroic (DMLP505L, Thorlabs) to separate the LED excitation and fluorescent emission.

To jointly record one-photon excitation and two-photon excitation, a 50:50 (reflectance: transmission) non-polarizing plate beam splitter (BSW27, Thorlabs) was placed after the second dichroic mirror to separate fluorescent signals for PMT (PMT1001, Thorlabs) and Airscope imaging system, respectively. Two fluorescence filters (MF525-39, Thorlabs) were configured in front of both the PMT and the Airscope to fully block both femtosecond laser and LED excitation beam. The back aperture of the objective was optically conjugated to the detection surface of the PMT with a 4f system (TTL200-A and AC254-050-A, Thorlabs).

To avoid excitation crosstalk and protect PMT from the Airscope path emission photons, we added a liquid-crystal electronically controlled shutter (LCC1620A/M Thorlabs) for the PMT detection path, which blocked the Airscope fluorescent photons when LED excitation was on. We further configured the EOM to be blocked during Airscope imaging. The LED was configured in trigger mode with a typical rising and falling time less than 1 ms, with further reduced duration time to avoid PMT overexposure. Both two-photon and Airscope captures were at 10 fps.

Using this hybrid system, we first aligned the focal planes of the two modalities with fluorescent beads and registered the fields of view of the two-photon and Airscope systems. We then performed temporally interleaved two-photon and Airscope imaging over the same regions in mice. Neurons extracted from the two-photon recordings using Suite2p²⁵ with additional manual labeling were treated as the reference ground truth. Meanwhile, Airscope data were processed using the DeepWonder pipeline with the same parameter settings, including background removal and neuronal signal extraction and demixing.

To evaluate the accuracy of Airscope recording, we compared the neurons identified from Airscope recordings with the reference neurons extracted from the corresponding two-photon

recordings, and quantified performance using F1 score and recall (mean F1 score = 0.78 ± 0.05 , mean recall = 0.81 ± 0.08 , $n = 36$ recording sessions). In addition, for all matched neurons, the extracted activity traces showed strong agreement between the two modalities, with a mean Pearson correlation of 0.82 ± 0.14 ($n = 1,196$ neurons). We also present raw data of Airscope, processed data using the DeepWonder pipeline, and reference two-photon ground truth images, which demonstrate good consistency (**Fig. R17**).

By comparing the results with the gold-standard two-photon imaging data, we found strong agreement between the two modalities. This demonstrates that the Airscope system is capable of providing reliable single-cell resolution in the mouse cortex, where neurons are densely packed. These extensive verifications have been incorporated into the second Results section of the revised manuscript.

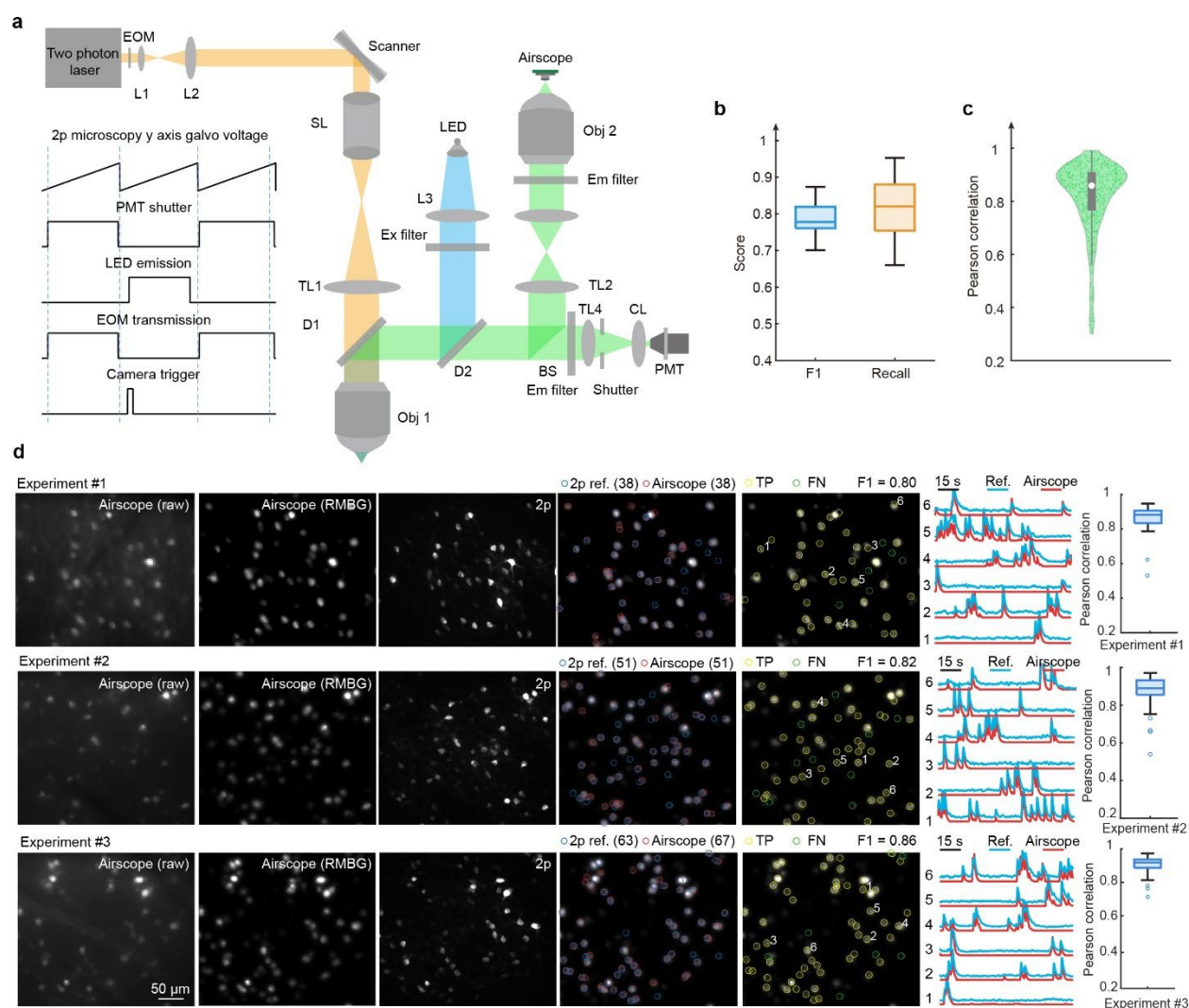


Fig. R17. Experiment validation of Airscope with synchronized two-photon imaging. **a**, Schematic of the Airscope and local two-photon imaging system used for validation. EOM, electro-optical modulator; L, lens; TL, Tube lens; M, mirror; D, dichroic mirror; BS, beam splitter; Em filter, emission filter; Ex filter, excitation filter; D1, dichroic (DMLP650L); D2, dichroic (DMLP505L,

Thorlabs); Shutter, liquid-crystal electronically controlled shutter; PMT, Photomultiplier tube. **b**, Boxplots of F1 score and recall for neuron detection by Airscope, compared with two-photon ground truth data across 3 animals and 36 recording sessions. (Central line: median; box: interquartile range; whiskers: maximum and minimum). **c**, Pearson correlation of temporal activity between Airscope and two-photon recordings ($n = 1,196$ neurons). (White circle: median; thick gray line: interquartile range; thin lines: upper and lower values; transparent disks: data points; transparent violin: data distribution). **d**, Evaluation of neuronal imaging fidelity of Airscope with functional ground truth obtained by two-photon microscope. Raw captured Airscope images, background-removed using the DeepWonder pipeline results and two-photon images captured simultaneously are presented in the first three columns. The fourth column exhibits overlaid neuron segments in two-photon (blue) and Airscope (red). The fifth column exhibits segments of true positives (TP, yellow), false negatives (FN, green). The sixth column presents randomly selected calcium traces by two-photon microscope (blue) and Airscope (red). The two-photon traces are intentionally shifted for clarity. The seventh column presents boxplots of correlation scores between two-photon microscope and Airscope. Central line inside the box: median. Box: interquartile range. Whiskers: maximum and minimum. Outliers: individual data points. $n = 3$ recordings across 3 brain positions are presented in 3 rows.

Major Comment 4 — Ground truth for neuron counts and imaging depth

4. The claim of the number of neurons imaged seems somewhat difficult to reliably measure with 1P excitation. The claims seem to be based on how many soma-like shapes can be detected but not based on the reality of how many neurons are present in the tissue, or even how many somas are labelled vs not labelled, i.e. the density of labelling. Given that the techniques used to label the tissue what fraction of the total number of neurons in the area did you image, only the brightly labelled neurons? How was it determined that cortical layer 2/3 are being imaged? Please provide some ground-truth to the imaging claims of number of neurons imaged in each of the experiments.

Response: We thank the reviewer for this important question, which we address in detail below.

1. Neuron detection methodology. We wish to clarify that our neuron identification pipeline (DeepWonder) does not rely solely on soma-like morphology. The algorithm evaluates candidates based on multiple criteria including temporal firing properties, and spatiotemporal dynamics consistency.

Briefly, the DeepWonder network estimates foreground neuronal information and fluorescence background, subsequently isolating them to generate neuronal ROI candidates. To be classified as a genuine neuron, a candidate must satisfy a suite of rigorous criteria: it must exhibit characteristic neuronal morphology, demonstrate calcium transient kinetics consistent with physiological activity (defined by specific decay constants and signal-to-noise ratio thresholds), and pass spatial compactness benchmarks. These criteria align with the validation protocols previously detailed in our response to **Reviewer 1, Blocking Comment 3**.

We have added a new Supplementary Note 2, “Neuronal signal processing,” to better describe the calcium signal processing pipeline.

2. Ground truth for neuron. To further validate the accuracy of detected neurons, we constructed a hybrid system to acquire interleaved synchronized Airscope and two-photon data, as detailed in our response to Major Comment 3. This allowed us to obtain a mean F1 score of 0.78 ± 0.05

cortical neuronal density is $\sim 10^5$ neurons/mm² across layers ^{26,27}, the results from tissue section staining, the proportion of neurons expressing fluorescence is estimated to be between 10% and 40%, though this value is dependent on the concentration of the labeling agent Tamoxifen used, given the sparse labeling characteristic of this line, Airscope detects a fraction comparable to that reported by established benchtop mesoscopic systems using identical transgenic lines (e.g., Sadegh Ebrahimi et al., 2022). We have now included a quantitative comparison of detected neuron density (neurons/mm²) between Airscope and published tabletop mesoscope recordings (Response Table R4), confirming comparable neuronal detection density. We claimed that Airscope achieve comparable neuronal detection efficiency compared to other systems, with a neuronal density of approximately 300 neurons/mm².

Table R4. Number of neurons detected using 1p methods in Rasgrf-Ai148D transgenic line

Reference	Imaging method	Detected neurons	FOV (mm ²)	Density (#/mm ²)
Sadegh Ebrahimi et al.(2022) ²⁸	Wide field (1p excitation)	5292	16	~330
Yuanlong Zhang et al. (2024) ²⁹	Light field (1p excitation)	18000	48	~400
Airscope (This manuscript)	Wide field (1p excitation)	8500	28	~300

4. Imaging depth and layer identity. To confirm the imaging depth and cortical layer, as well as the range of neuronal expression, we first performed DAPI staining in a sagittal brain sections. Subsequently, we used two-photon imaging to visualize the expression of GCaMP6f. As shown in **Fig. R19**, we observed that GCaMP6f expression occurred between depths of 100-350 μ m, confirming that we were imaging the cortical layers 2/3. Expression was densest between 100-250 μ m, and became progressively sparser below 250 μ m.

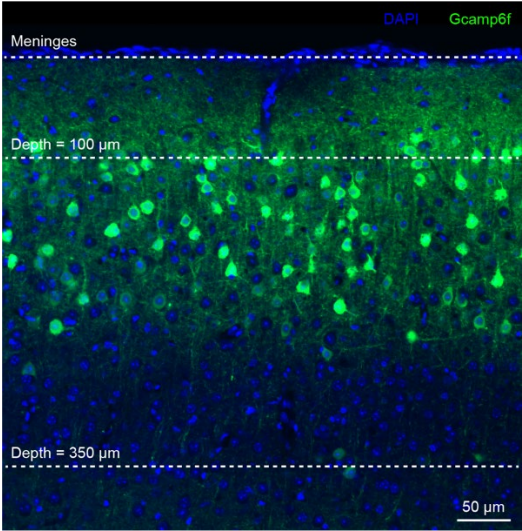


Fig. R19. Two-photon imaging of GCaMP6f expression in cortical layers of DAPI-stained brain section. DAPI (blue) and GCaMP6f (green) staining in a coronal brain section. Scale bar: 50 μm .

To further determine the imaging depth and the fidelity of the Airscope neuronal imaging results at different depths, we performed a hybrid system that simultaneously records neuronal signals using two-photon (2P) and Airscope imaging modalities across different depths beneath the cortical surface. The recordings were performed at multiple depths (100 μm , 150 μm and 200 μm , $n = 3$ mice), allowing us to assess the performance of Airscope across various regions of the brain (**Fig. R20**). Although the detection accuracy slightly decreased with increasing depth due to scattering, the F1 score were consistently maintained over 0.75, demonstrating a relatively high performance across all depths. Additionally, the Pearson correlation of trace activity remained approximately 0.8. These results suggest that Airscope is effective for imaging at various depths within 200 μm beneath the meninges, despite the challenges posed by increased tissue thickness.

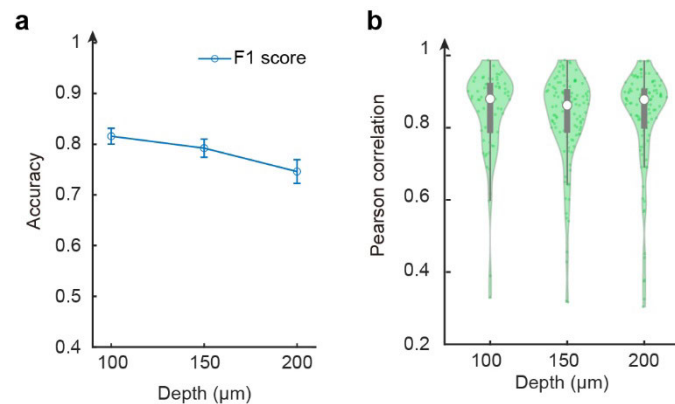


Fig. R20. Evaluation of neuronal detection by Airscope across different depths. **a**, Evaluation of neuron-detection accuracy (F1 scores) achieved by Airscope at different cortical depths. Circle: mean. Error bar: ± 1 S.D. **b**, Pearson correlation scores achieved by Airscope at different cortical depth. White circle: median. Thick gray vertical line: interquartile range. Thin vertical lines: upper and lower proximal values. Transparent disks: data points. Transparent violin-shaped areas: kernel density estimate of data distribution.

Major Comment 5 — Wireless vs. tethered comparison & causal controls

5. It is not clear throughout the manuscript what the benefits of the system are compared to tethered system, it would be important to show an experiment from the same brain-tissue where activity maps could be showed using the untethered microscope that was different to activity maps when using the tethered scope. As it stands, the main argument is that tether free is better because of the application rather than a finding. This is ok, but with the lack of single-cell resolution and the lack of controls, especially controls for linking the correlated activity to behavior, there are observations of neuronal activity but no causative controls. This makes the biological claims somewhat difficult to understand.

Response: We thank the reviewer for this constructive suggestion, which we fully agree strengthens the manuscript. In the following, we first address the reviewer's concern regarding

single-cell resolution, which was substantiated in our previous response. Subsequently, we present a newly designed experiment characterizing the divergent behavioral movement patterns of animals equipped with Airscope versus those utilizing tethered miniaturized microscopes. Following this behavioral analysis, we performed a comparative evaluation of neuronal firing patterns across both conditions at the single-cell level. In the second section, we detail the development of our causal control instrumentation.

1. Single-cell resolution. We respectfully note that Airscope achieves 4-6 μm lateral resolution, which is sufficient to resolve individual somata in layer 2/3 (typical diameter $\sim 10\text{--}15\ \mu\text{m}$), as rigorously demonstrated in our response to Reviewer 1, Blocking Comment 1 (refer to **Supplementary Fig. 8**). Furthermore, we performed simultaneous Airscope and two-photon imaging experiments to validate our system's performance against a gold-standard 2P reference. Our results indicate that Airscope faithfully recovers single-cell activity, achieving a mean detection F1-score of 0.78 and maintaining a mean temporal correlation exceeding 0.8 relative to the two-photon ground truth (refer to **Fig. R17**). Collectively, these spatial and functional validations substantiate Airscope's capability for single-cell resolution imaging.

2. Tethered vs. untethered comparison. We have now performed paired recordings in the same animals and cortical regions using both tethered (wired power, wired data) and fully wireless configurations.

In the first scenario, we compared locomotor behavior in the same animals wearing either a untethered Airscope or a cable-tethered Airscope in a $40 \times 40\ \text{cm}^2$ openfield (**Fig. R21a**). The tethered configuration used a Cooner wire (CW8251), consistent with the head-mounted weight used in **Extended Data Fig. 2** and commonly adopted in tethered miniature microscopes such as Miniscope V3 and Miniscope V4. In this planar open-field arena ($40 \times 40\ \text{cm}^2$), the two configurations exhibited comparable locomotor behavior, including cumulative distance, speed distribution, total distance traveled, and 75th percentile speed (**Fig. R21c–f**). These results indicate that, in planar movement, tethering of a miniscope in animal's head will not cause large differences compared to fully untethered mode.

In the second scenario, animals were recorded in a $60 \times 60\ \text{cm}^2$ arena containing hexagonal pillars that form a heterogeneous 3D terrain (**Fig. R21g**). An RGB-D camera was used to capture vertical motion, and SAM2Mice enabled precise spatiotemporal detection of height-change events, including climbing and descending transitions. In contrast to the planar condition, tethered animals in the 3D environment exhibited altered locomotor patterns, including reduced cumulative exploration distance and lower locomotor speed (**Fig. R21h–i**). The total spatial area explored was also markedly reduced (**Fig. R21m**). This reduction was accompanied by a decreased capacity for vertical transitions, as reflected by a lower frequency of height-change events in tethered animals compared to the untethered condition (**Fig. R22c, d**).

We subsequently investigated whether these behavioral discrepancies manifested as measurably distinct dorsal cortical activity patterns. To address this, we longitudinally registered neuronal coordinates within the same animals across tethered and untethered conditions, enabling direct comparative analysis of identical neuronal populations during both experimental setups (**Fig. R22e and Methods**). Our analysis was specifically time-locked to height-change events, which are first classified into two behavioral categories, ascend (lower to higher platform) and descend (higher to lower platform), based on the direction of vertical displacement.

To visualize the neural dynamics of these overlapping populations, we projected the population activity vectors corresponding to task-relevant events into a lower-dimensional state space using principal component analysis (PCA) ⁴. Briefly, to construct PCA trajectories, individual neuronal activity was aligned to the onset of ascending or descending events within a defined peri-event time window. Trial-averaged activity for each event type was computed per neuron, yielding a time-resolved response profile. These profiles were concatenated across events to generate a high-dimensional activity vector for each neuron. Subsequently, vectors encompassing all neurons across all animals were aggregated into a unified population matrix. Following mean-centering, PCA was applied to delineate a common low-dimensional neural state space that captured the dominant variance structures within the population data. The top principal components, accounting for the majority of the variance, were retained as a shared orthogonal basis. For each behavioral condition (e.g., tethered-ascending or untethered-descending trials), we computed the mean population activity and projected these time-evolving vectors into the PCA space utilizing the previously derived loading matrix. Connecting the projected data points across consecutive time bins yielded neural trajectories, with each continuous curve representing the temporal evolution of the collective dorsal cortical population state for a specific condition and behavior.

We discovered that neural trajectories corresponding to descending and ascending behaviors exhibited characteristic clockwise (CW) rotational dynamics. Notably, both conditions drove a segregation of neural activity subspaces even prior to the onset of the height change, indicating the presence of separable brain states preceding trial initiation (**Fig. R22g**). Furthermore, animals in the tethered condition displayed elongated neural trajectories compared to their untethered counterparts (**Fig. R22g**).

To test whether dorsal cortical populations encode ascending and descending movements, we trained a support vector machine (SVM) classifier on single-trial neural activity. Decoding accuracy increased around the onset of height change in both conditions, peaking several frames after the transition (**Fig. R22h**). However, cross-condition transfer was minimal: decoders trained on tethered recordings performed at near-chance levels when tested on untethered data, and similarly, decoders trained on untethered recordings showed near-chance performance on tethered data (**Fig. R22h**). These results indicate that cortical activities of vertical movement are largely condition-specific, with limited generalization across tethered and untethered configurations.

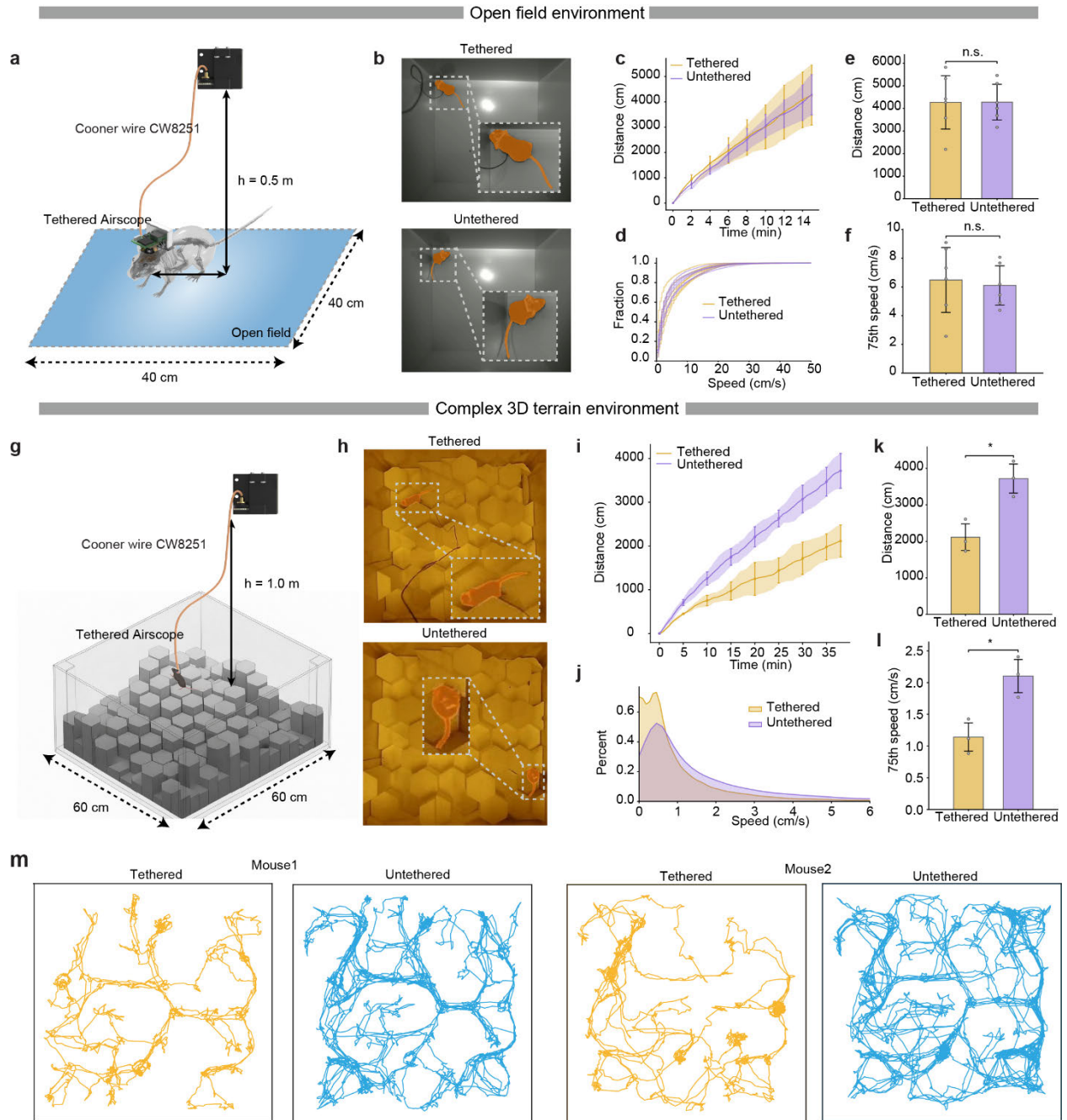


Fig. R21. Evaluation of movement agility under tethered and untethered Airscope configurations in open-field and complex 3D terrain assays. **a**, Schematic of the tethered Airscope configuration in a standard open-field arena (40×40 cm), with the device connected through a Cooner wire (CW8251) suspended at a height of 0.5 m. **b**, Representative images of mice in the open-field arena under tethered and untethered conditions, with overlaid masks segmented by SAM2Mice. **c**, Cumulative distance plot over 15 minutes; center lines indicate the mean and shaded areas indicate SD, $n = 5$ animals. **d**, Cumulative speed distribution in the open-field arena; each line indicates one individual animal. **e**, Bar plot of total distance traveled across tethered and untethered groups (two-sided Wilcoxon rank-sum test, $p = 0.985$, $n = 5$ animals). Bars indicate mean, error

bars denote SD, and dots represent individual animals. **f**, Bar plot of seventy-fifth percentile speed between tethered and untethered conditions (two-sided Wilcoxon rank-sum test, $p = 0.761$, $n = 5$ animals). Symbol same as **e**. **g**, Schematic of the tethered Airscope configuration in a complex 3D terrain arena (60×60 cm), with the device connected through a Cooner wire (CW8251) suspended at a height of 1.0 m. **h**, Representative images of mice navigating the complex 3D terrain under tethered and untethered conditions, with overlaid masks segmented by SAM2Mice. **i**, Cumulative distance over time in the complex 3D terrain; center lines indicate the mean and shaded areas indicate SD. **j**, Speed distribution in the complex 3D terrain across tethered and untethered conditions. **k**, Bar plot of total distance traveled in the complex 3D terrain across tethered and untethered groups (two-sided Wilcoxon rank-sum test, $*p = 0.0366$, $n = 3$ animals). Symbol same as **e**. **l**, Bar plot of seventy-fifth percentile speed in the complex 3D terrain (two-sided Wilcoxon rank-sum test, $*p = 0.0309$, $n = 3$ animals). Symbol same as **e**. **m**, Representative locomotor trajectories from two mice in the complex 3D terrain under tethered (orange) and untethered (blue) conditions.

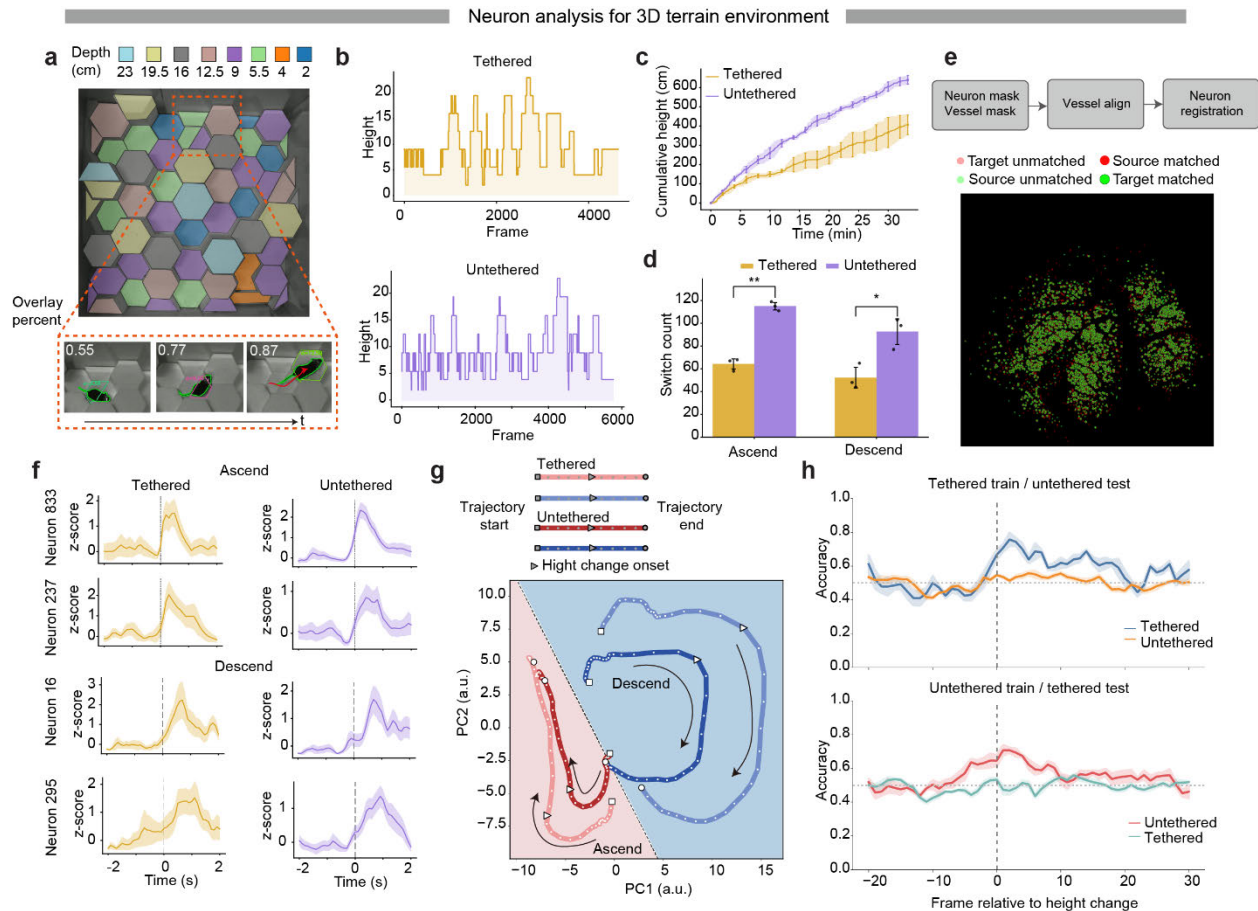


Fig. R22. Neuronal analysis of animal wearing tethered and untethered miniaturized microscope. **a**, Terrain height annotation and readout. Individual terrain tiles are annotated and measured, colored masks overlaid indicate different height (23, 19.5, 16, 12.5, 9, 5.5, 4, and 2 cm). The mouse is segmented using SAM2Mice, and height is assigned from the terrain region occupied by the

mask. Representative segmentation examples are shown below. **b**, Example height traces across time for the tethered and untethered configurations. **c**, Cumulative height change over time in the complex 3D terrain across tethered and untethered groups; center lines indicate the mean and shaded areas indicate SD. **d**, Bar plots of detected ascend and descend transitions in the two recording configurations (two-sided Wilcoxon rank-sum test, $**p = 0.004$ for ascend transitions and $*p = 0.0185$ for descend transitions, $n = 3$ animals). Height changes exceeding 4 cm are classified as ascend or descend events. Bars indicate mean, error bars denote SD, and dots represent individual animals. **e**, Cross-session neuron registration between recording configurations. Neuron masks from the untethered dataset (source, $n = 3,143$ neurons) and the tethered dataset (target, $n = 2,567$ neurons) were aligned by vessel landmarks, and 1,534 matched neurons are identified. **f**, Activity tuning of the same registered neurons during ascend and descend behaviors (Untethered: descend, $n = 24$ trials; ascend, $n = 35$ trials. Tethered: descend, $n = 14$ trials; ascend, $n = 15$ trials. Center line: mean, shaded area: SEM). **g**, Population neural trajectories projected into a lower-dimensional common principal component (PC) subspace based on the matched population ($n = 3$ animals, 6682 neurons in total; total trials: untethered descend = 98, untethered ascend = 114, tethered descend = 66, tethered ascend = 78). **h**, Decoding of ascend versus descend states using a support vector machine (SVM) classifier. Solid lines indicate mean and shaded areas indicate SEM across five-fold validation. Top: Classifier trained on tethered condition and tested on untethered recordings. Bottom: Classifier trained on untethered condition and tested on tethered recordings.

We have now added the **Fig. R21** and **Fig. R22** as new **Extended Data Fig. 3**. In the revised manuscript, we have described the experiment in the second results section; it reads

"To evaluate the behavioral range afforded by untethered recording, we compared locomotor agility in the same animals wearing the Airscope or a tethered miniscope across 2D and 3D environments. The two configurations were comparable in 2D open-field exploration but diverged in a 3D arena, where Airscope-equipped animals showed greater agility and more height-change events (Extended Data Fig. 3). Tracking the same neuronal populations across conditions revealed condition-specific dorsal cortical dynamics that were consistent across mice: PCA trajectories were more compact under the Airscope, and an SVM decoder trained on one condition generalized poorly to the other (Supplementary Note 3). The Airscope thus matches tethered performance in 2D while accessing a broader behavioral and different cortical activity patterns in 3D."

3. Optogenetic perturbation controls. To address questions of causality within cortical networks, in this revision, we engineered a dual-LED optogenetic inhibition module compatible with the Airscope platform.

As shown in **Fig. R23**, the newly developed optogenetic Airscope features an integrated array of six light-emitting diodes (LEDs), comprising three LEDs dedicated to fluorescence imaging (470 nm) and three allocated for optogenetic manipulation (570 nm, targeting Chrimson). The imaging and optogenetic stimulation channels are independently programmable, as shown in **Fig. R23**. Consequently, when coupled with real-time behavioral or neural state-decoding algorithms, this platform provides the necessary hardware foundation for closed-loop optogenetic interrogation.

Although the current investigation focuses on characterizing state-dependent population dynamics rather than dissecting directed functional connectivity, the development of this optogenetic prototype establishes a critical technological foundation for future interrogations of brain-wide causal pathways. Notably, to the best of our knowledge, no comparable simultaneously mesoscale

optogenetic and imaging device currently exists for freely moving animals, particularly in a fully wireless configuration.

[Figure Redacted]

Fig. R23. Modification of Airscope for optogenetics capability. **a**, 3D schematic of the optogenetics-enabled Airscope architecture. **b**, Assembled device shown alongside with a 19-mm coin for scale. **c**, Custom FPC-based six-LED array. **d**, Emission spectra of the blue LED for imaging excitation and the yellow LED for optogenetic stimulation. **e**, Independent control of the two LED channels in four operation modes: imaging plus stimulation, imaging only, stimulation only, and all off.

Major Comment 6 — Overclaiming in biological interpretations

6. There are many examples of overclaiming for cortical function in which the data does not stand up. I will give a few examples, but these are not by any means the only examples, please go through the manuscript and remove or rewrite these to fit the data accurately.

Response: We sincerely thank the reviewer for this careful reading. We fully agree that precise language is essential, and we have conducted a thorough revision of the entire manuscript to ensure all claims are commensurate with the supporting data. Below we address each specific example:

6a — Abstract: "initial appraisal of untethered brain-behavior dynamics"

From the abstract, "39 Using Airscope, we provide an in 40 initial appraisal of untethered brain-behavior dynamics. I' m not sure what an initial appraisal is in this context, especially in the abstract."

Response: We agree this phrasing was imprecise. We have revised the abstract to read: " *Using Airscope, we demonstrate versatile applications of untethered cortical imaging across diverse behavioral paradigms.*"

6b — Abstract: "reward discovery state"

42 "comes across animals. We identify a previously uncharacterized reward discovery state 43 during maze exploration." I cannot find the data and control experiments that literally show this finding beyond correlation.

Response: We appreciate this feedback. We have revised the statement to: "*We observe reward-location-associated cortical activity patterns during maze exploration.*"

To further address this concern, we added new control analyses in **Fig. R24** to test whether reward-related activity could be explained by simple neuron–reward correlations or conventional linear decoding approaches. First, we examined Pearson correlations between individual-neuron calcium activity and reward-state labels. Correlation values were centered near zero and were comparable to shuffled controls (**Fig. R24b, c**), indicating that single neurons did not exhibit strong or consistent reward tuning. Session-averaged correlations were not significantly different between observed and shuffled data (two-sided Wilcoxon rank-sum test, $p = 0.878$, $n = 3$ animals, 3 sessions per animal). We next tested whether reward states could be recovered from population activity using a standard linear decoder. Linear discriminant analysis (LDA) projections showed no clear separation between reward and non-reward frames in either observed or shuffled datasets (**Fig. R24d, e**). Consistently, within-cluster and between-cluster distances were not significantly different for either observed data ($p = 0.725$) or shuffled controls ($p = 0.940$; two-sided Wilcoxon rank-sum tests). Furthermore, 3-fold cross-validation accuracy of the LDA decoder remained close to chance level and did not differ from shuffled controls ($p = 0.5$, two-sided Wilcoxon rank-sum tests; **Fig. R24f**). This analysis follows established strategies in systems neuroscience for dissociating behavioral variables from neural population activity, as exemplified by Musall et al., 2019³⁰ and Stringer et al., 2019³¹.

In contrast, in **Fig. 5j**, rSLDS identified a distinct HMM state associated with reward position, suggesting that reward-related structure is expressed not as a simple single-neuron correlation or linearly decodable population signal, but as a latent dynamical state. Nevertheless, we agree that the original phrasing overstated this finding, and we have revised the manuscript accordingly. We also add these control experiments into labyrinth exploration section, it now reads "*Simple correlation and linear decoding by LDA revealed no significant association between dorsal cortical activity and reward location, indicating this reward-related structure was not trivially explained by activity–reward correlations.*"

We have now added **Fig. R24** to Supplementary Fig. 15.

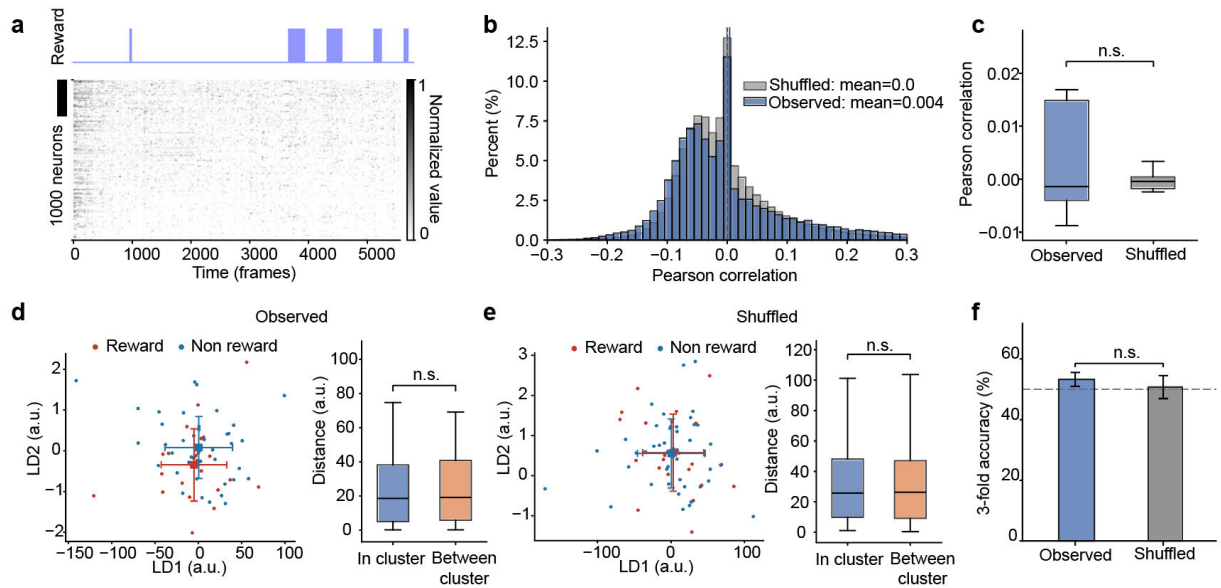


Fig. R24. Reward location is not captured by single-neuron activity or conventional linear decoding. **a**, Example calcium activity heatmap during maze exploration, with reward-associated frames marked in blue. **b**, Histogram of Pearson correlations between individual-neuron activity and reward state for observed and shuffled data. Both distributions are centred near zero. **c**, Boxplot of average Pearson correlation for recorded sessions between observed and shuffled controls (two-sided Wilcoxon rank-sum test, $p = 0.878$, $n = 3$ animals, 3 sessions per animal). n.s., not significant. **d**, Left, LDA projection of population activity in observed data for validation set, with reward frames shown in red and non-reward frames shown in blue. Right, boxplot of within-cluster and between-cluster distances (two-sided Wilcoxon rank-sum test, $p = 0.725$, $n = 3$ animals, 3 sessions per animal). Centre line, median; box, interquartile range; whiskers, data range. **e**, Same analysis as **d** for shuffled data (two-sided Wilcoxon rank-sum test, $p = 0.940$, $n = 3$ animals, 3 sessions per animal). **f**, 3-fold cross validation accuracy of LDA for distinguishing reward from non-reward frames in observed and shuffled data (two-sided Wilcoxon rank-sum test, $p = 0.5$). n.s., not significant. Bar plots represent mean $\pm$ SD.

6c — Abstract: "causal pathway from spinal cord to cortex"

43 We map a new causal pathway from spinal cord to the cortex 44 causal flow during fear conditioning enabled by simultaneous brain-spinal recordings. I cannot find data that shows a causal pathway. This implies anatomy and causal experiments

Response: We thank the reviewer for this correction and agree that "causal pathway" overstated our claim. We have revised the statement to: "characterize Granger-causal influence from spinal cord to cortex during fear conditioning, enabled by simultaneous brain-spinal recordings." Our Granger causality analysis is causal in the Wiener–Granger sense: it establishes that past spinal activity carries predictive information about future cortical activity beyond what cortical history alone explains, thereby quantifying directed statistical dependence and temporal precedence of information flow. It does not, on its own, establish anatomical connectivity or interventional

causation, and we cannot exclude contributions from unobserved common drivers or indirect polysynaptic routes from our recordings alone.

To provide complementary perturbation evidence for the spinal-to-cortical direction of this influence, we performed spinal cord transection experiments (**Fig. R25**). As illustrated in the experimental schematic (**Fig. R25a**), hindpaw electrical stimulation was delivered either in control animals or following complete spinal cord transection, which physically disrupts ascending spinal-to-cortical transmission. Control animal shows stimulus-locked cortical activity, whereas this response is largely absent after spinal cord transection (**Fig. R25b**).

Following spinal cord transection, individual neuron that previously responded to the electrical stimulation now exhibited substantially attenuated evoked responses to the same stimulus (**Fig. R25c**). Furthermore, across the board cortical regions captured by Airscope, neuronal populations that initially displayed widespread stimulus-evoked activity were profoundly suppressed post-transection (**Fig. R25d**). Quantitatively, the trial-averaged single-neuron response dropped from ~ 0.21 to near zero (**Fig. R25e**; paired t-test, **** $p = 2.052 \times 10^{-213}$), and the fraction of stimulus-responsive cortical neurons fell significantly from $\sim 22\%$ to $<1\%$ across three animals (**Fig. R25f**; paired t-test, * $p = 0.037$). Together, these perturbation results confirm that ascending spinal input is necessary for driving the observed cortical activity, providing interventional support for the directed spinal to cortical influence characterized by our Granger causality analysis.

We also corrected all "casual analysis" into "Granger-causality analysis" in the revised manuscript.

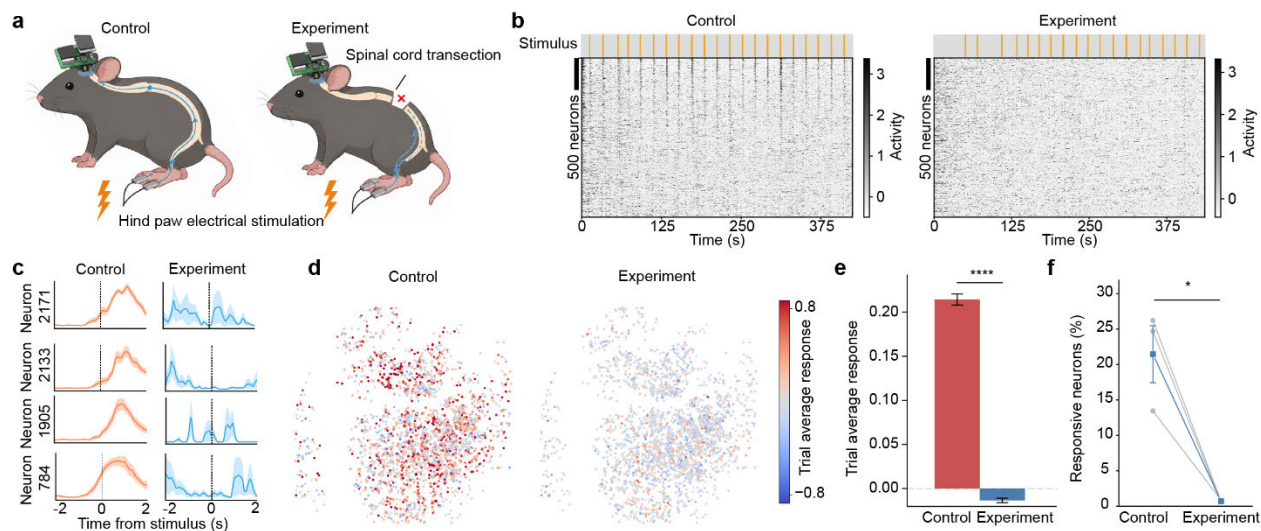


Fig. R25. Spinal cord transection abolishes cortical responses evoked by hind-paw stimulation. **a**, Experimental design. For control groups, hind-paw electrical stimulation is delivered during Airscope calcium imaging. For perturbation groups, the spinal cord is transected prior to stimulation to disrupt spinal to cortex transmission. **b**, Population neuron activity heatmaps aligned to repeated hindpaw stimulation epochs (orange lines) for control and spinal-transected mice. Neurons are matched across sessions by neuron registration and sorted according to their trial average responses. **c**, Example stimulus-aligned calcium responses from the same registered neurons shown before and after spinal transection. Center lines indicate mean and shaded areas indicate SEM across trials. Dashed lines indicate stimulus onset. **d**, Spatial distribution of trial-

averaged stimulus responses from the same imaging field in control and transected conditions. Each dot represents one neuron, color-coded by trial-averaged response amplitude. **e**, Bar plot of single neuron trial-averaged response in control and spinal-transected mice (paired t-test, **** $p = 2.052 \times 10^{-213}$; $n = 3$ mice, 4772 registered neurons in total; mean $\pm$ SEM). **f**, Fraction of stimulus-responsive cortical neurons in each animal before and after spinal transection (paired t-test, * $p = 0.0369$). Grey lines indicate individual animals; blue error bars indicate mean $\pm$ SEM.

6d —Introduction: "Airscope unveiled cortical basis of animals' natural exploration in a labyrinth"

437 Airscope unveiled cortical basis of animals' natural exploration in a labyrinth How can this be true? If only layer 2/3 were being imaged, then what is the rest of the cortex doing during this task? Where are the control experiments? Where is the raw data that shows this effect? How does the model work? Saying that the cortical basis of an animal's maze exploration has been unveiled is literally not true

Response: We thank the reviewer for flagging this overstatement. The reviewer is correct: Airscope images layer 2/3 across dorsal cortex, and from these data alone we cannot make a comprehensive claim about the "cortical basis" of maze exploration. We have substantially softened the framing of this section so that its purpose is no longer to *explain* maze behavior, but to demonstrate (i) that Airscope's untethered, large-FOV imaging makes it experimentally feasible to record cortex-wide layer 2/3 population activity during freely-moving navigation in a complex environment, and (ii) that, within these data, population-level latent dynamics, rather than single-neuron turn-direction coding, dominate the explainable variance.

In the revised manuscript, we have revised the section title into "*Airscope enables large-scale cortical population imaging during free labyrinth navigation*". We have also removed the phrase "unveiled cortical basis" everywhere it appeared in the main text, figure legends, and Abstract, and replaced it with language describing experimental capability and population-level observations.

Major Comment 7 — Open-source release of full designs

7) Open-source access to these technologies would maximize the impact of this advance on the field. But to be open source, the authors need to fully release electronic and optical designs. Crucially, Zemax designs must be made available for future improvements of the system and to give scientists that might use the system in the future the full understanding of the data generated with this device.

Response: We fully agree with the reviewer and appreciate this emphasis on open science. In this revision, we have released the complete design package, including:

- Full Zemax optical design files (.zmx) with all lens prescriptions and optimization merit functions
- Electronic schematics and PCB layout files
- Mechanical CAD files for all housing components
- Firmware source code for wireless transmission
- Assembly instructions and alignment procedures

All materials are available at our project repository (URL) and referenced in the Data Availability section of the manuscript.

Major Comment 8 — Weak discussion and missing references

8) The discussion is weak and missing many many key references, especially in light of the head-mounted technology and biological claims in the manuscript. It would be important to discuss the microscope in terms of the other head-mounted microscopes that employ 3 and 2-photon excitation and provide optical sectioning. All microscopes have both pros and cons over the current microscope, these should be discussed succinctly and in a balanced approach. In another example, the current method attempts to scene reconstruction by adding another camera to the head, this has been done before and these papers should be cited. Scene reconstruction of freely moving mice has been achieved by two methods, one by a head mounted camera (Parker et al 2022 Joint coding of visual input and eye/head position in V1 of freely moving mice) and by eye-tracking and modeling (Holmgren 2020 Visual pursuit behavior in mice maintains the pursued prey on the retinal region with least optic flow)

Response: We thank the reviewer for this important critique. We agree the original Discussion was too narrow and under-cited the broader landscape of head-mounted imaging. We have substantially rewritten it and add a paragraph to provide a balanced, pro/con comparison with 2P and 3P head-mounted systems that explicitly acknowledges their optical-sectioning advantage. The full revised Discussion is reproduced below:

Airscope complements, rather than competes with, existing head-mounted multiphoton microscopes. Head-mounted two-photon (2P) systems provide optical sectioning, reduced out-of-focus background, subcellular and multiplane imaging, and access to deeper or implanted targets, with recent implementations extending these capabilities to larger populations and greater depths (Helmchen et al., 2001; Zong et al., 2017, 2021, 2022). Head-mounted three-photon (3P) microscopes push sectioned imaging further, enabling recordings across all cortical layers and, in one implementation, dorsal hippocampal CA1, albeit at reduced imaging speed and under tighter constraints on excitation power and tissue exposure (Klioutchnikov et al., 2020, 2023; Zhao et al., 2023). Airscope makes a complementary trade-off: it forgoes optical sectioning and deep-layer access in exchange for low head-borne weight, wide cortical coverage, simpler optics, and (critically) fully untethered, wireless operation. Current head-mounted 2P and 3P microscopes rely on hollow-core fibers for femtosecond pulse delivery, which has so far precluded wireless operation; Airscope's single-photon design sidesteps this constraint and is therefore well suited to naturalistic behaviors that require unrestricted movement.

We also place our simultaneous brain and facial expression recordings (original **Fig. 6l-p**) and concurrent brain and environmental recordings (original **Fig. 6q-u**) into “Supplementary Note 5 | Airscope enables multimodal observation beyond neuronal imaging”, to maintain the manuscript's focus on the flagship applications. We noticed Meyer et al. (2018) established lightweight head-mounted camera monitoring in freely moving mice; Parker et al. (2022) combined eye/head measurements with visual input from the mouse's perspective; and Holmgren et al. (2021) reconstructed the visual field from eye tracking and geometric modeling. Thus, the advance here is not scene reconstruction per se, but its co-integration with a 1-g mesoscope for simultaneous cortex-scale cellular imaging and scene capture in a fully untethered preparation. All these publications are properly cited (see **Supplementary Note 5**).

Minor aspects:

Minor Comment 1 — Resolution degradation toward edges

Authors report uniform 4 μm resolution, but it does seem to degrade on the MTFs towards the edges? The authors extensively quantify the resolution and report correctly the values in the result paragraph (1.219-232), but later only mention central resolution, while it clearly degrades towards periphery (ex. 1.244, table1 only 4 μm value is reported). Upper bound or range should be used instead.

Response: We thank the reviewer for this careful observation. We agree that reporting only the central resolution was incomplete.

As shown in the response to Blocking comment 1 of Reviewer 1, this time we have conducted a more detailed FOV-aware resolution plot measured by 1 μm point emitters., as shown in **Fig. R1**. Specifically, the 6 mm FOV was divided into four concentric annular regions according to lateral displacement from the FOV center. Within each annular region, six 1 μm point emitters were analyzed to statistically quantify the spatial dependence of the lateral resolution. The mean lateral resolutions measured in the four regions were **3.99 μm , 4.02 μm , 4.82 μm , and 6.04 μm** , respectively, corresponding to increasing lateral displacement from the center to the edge of the FOV. This annular-region-based analysis provides a more accurate characterization of the gradual resolution degradation from the center toward the periphery.

Thus, in the revised manuscript, **Supplementary Table 1** has been updated to include both central and peripheral resolution values, and all text references now state the full range: 4–6 μm lateral resolution across the 6 mm FOV.

Minor Comment 2 — "50 $\times$ improvement" claim is misleading

L623-624 “optical design that leverages aspherical lens systems to improve per gram space - bandwidth product over 50 times” This statement is confusing. By comparing to optics in the closest competitor, that also used aspheric lens system the improvement factor estimated by the authors is 2 and normalized by weight 2.8 (although this is weight of full system, not just optics). Factor 50 compares with closest wireless competitor, but that is irrelevant to optics.

Response: We thank the reviewer for identifying this ambiguity. We respectfully note that the Kiloscope, while optically competitive, relies on fiber-coupled illumination and tethered data transmission, which fundamentally precludes untethered operation — the central design objective of Airscope. A detailed comparison of Airscope with the Kiloscope has been provided in **Major Comments 1 and 2** of Reviewer 2.

In the revision, we have revised the statement to clearly specify the comparison basis. Additionally, the improvement factor in per-gram space-bandwidth product has been updated following a careful recalculation that accounts for the resolution drop at the margins of the field of view (see **Fig. R1** and **Fig. R2**). After incorporating this correction, we recalibrated the improvement factor to approximately 40 $\times$ (compared to Miniscope Wirefree, Wireless Miniscope, and Finchscope, see **Fig. R26**). Detailed resolution, FOV, throughput, and bandwidth comparison can be seen in **Supplementary Table 1**.

Nevertheless, following the instruction of the reviewer, we revised the manuscript to explicitly mention the comparison comes only with wireless counter part. The revised Discussion text now reads: " *An aspherical lens system improves per-gram space-bandwidth product ~40-fold to previous wireless system, and systematically optimized illumination, sensors, data compression, and wireless transmission eliminate the constraints of tethered recording*"

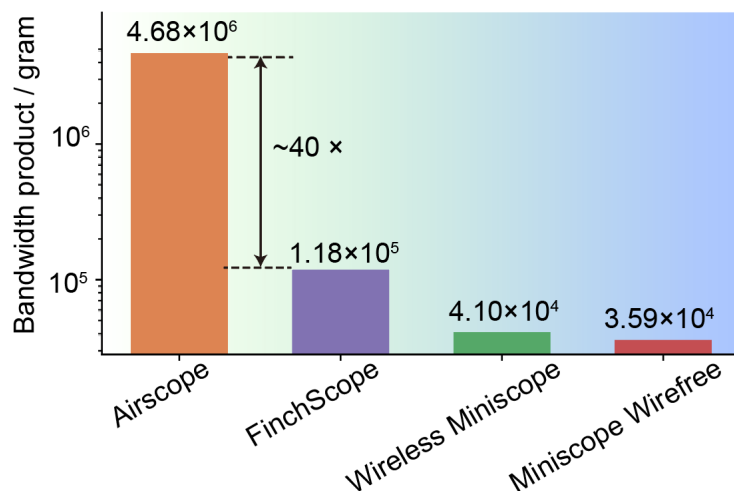


Fig. R26. Space-bandwidth product per gram compared to state-of-the-art wireless microscopes.

Minor Comment 3 — Contribution centered on transmission.

The true outstanding advance here is the wireless transmission, that maintains sufficiently high data rate and is implemented in a lightweight and compact footprint.

Response: We appreciate the reviewer’s perspective; however, we respectfully maintain our original position. The development and integration of miniaturized microscopes necessitate a strategic balance across multiple performance metrics, a balance successfully achieved by Airscope. Specifically, regarding wireless functionality, Airscope offers integrated capabilities that are absent in the previous architecture.

The technical advantages of Airscope are comprehensive, encompassing:

- **Optical Engineering:** Implementation of a small-form-factor lens system for mesoscale neuronal recording with high performance.
- **Data Transmission:** High-fidelity, untethered wireless telemetry.
- **Systems Architecture:** A ruggedized, highly compact design optimized for durability and portability during unrestrained behavior. This design is also advocated by Reviewer #1 (see non-blocking comment, 5c).

Furthermore, the Airscope framework is inherently scalable. To empirically demonstrate this modularity, we have expanded the platform's capabilities as detailed in our revised manuscript:

- *Response to Reviewer 2, Major Comment 1:* We leveraged our curriculum learning design paradigm to engineer two distinct optical modules: a centimeter-scale lens system

maintaining single-neuron resolution, and an ultra-compact system featuring a 1 mm field-of-view (FOV) with 1 μ m spatial resolution.

- *Response to Reviewer 2, Major Comment 5:* We designed a fully integrated optogenetic Airscope. This prototype comprises three LEDs dedicated to fluorescence imaging (470 nm) and three independent LEDs allocated for optogenetic manipulation (570 nm).
- *Response to Reviewer 2, Minor Comment 4:* We further integrated micro-LED-based optogenetics to enable simultaneous cortical imaging and targeted subcortical neuromodulation. Utilizing this interface, we successfully demonstrated a simultaneous imaging and seizure management protocol.

We therefore contend that the Airscope platform represents a transformative, highly versatile advancement for neural imaging in freely moving animals.

Minor Comment 4 — Lack of a clear enabling experiment

The authors show extensive application possibilities enabled by the absence of wires in a truly impressive set of behavioral paradigms. But it somehow lacks one clear example, where this key innovation giving access to completely unrestrained behavior enables new types of scientific questions. I do not mean here a niche experiment paradigm, like swimming or sheltering, but rather experiments enabling novel approaches to general neuroscience questions.

Response: We thank the reviewer for this insightful comment. We agree that a compelling demonstration of scientific questions uniquely enabled by untethered imaging strengthens the manuscript significantly.

1. Tethering distorts three-dimensional behavior and cortical population geometry — a comparison only wireless imaging makes possible.

As we previously responded to **Major Comment 5, Reviewer 2**, we conducted experiments that clearly compare tethered and untethered animals in both behavior patterns and cortical population activity. To assess whether tethering introduces behavioral or neural artefacts in head-mounted imaging, we performed paired recordings in the same animals and cortical regions under wired and fully wireless configurations. In 2D open-field sessions, tethered and untethered animals were indistinguishable on every gross locomotor metric tested, including cumulative distance, the cumulative distribution of speed, total distance and the 75th-percentile speed, indicating that planar movement is largely insensitive to cable load. The tether did, however, impose clear constraints once the task demanded volumetric movement in a 60 $\times$ 60 cm arena populated with hexagonal pillars. Behavior analysis pointed out tethered animals travelled shorter distances at lower speeds (**Fig. R21h–l**), covered a smaller fraction of the arena (**Fig. R21m**) and performed fewer height-change events (**Fig. R22c, d**).

Because neurons were registered longitudinally across conditions, we could ask whether this behavioral restriction was mirrored in cortical activity. PCA analysis unveiled that the activities of the same neuron population were already separable before movement onset, indicating preparatory brain states, and tethered trajectories were consistently elongated relative to untethered ones. An SVM decoder trained on one condition showed weak transferability to the other, indicating that the cortical representations underlying vertical movement differ substantially between tethered and untethered conditions.

Together, these results show that tethering leaves planar locomotion intact but distorts 3D behavior and the cortical population geometry that supports it, and that Airscope recovers both.

2. Sharpen articulation of the scientific question that the wireless system enables.

In this revision we have refocused and substantially expanded the multi-animal social interaction experiment around such a question is *competitive rank reflected in distributed cortical population activity beyond prefrontal cortex, and does the inter-brain neural coupling reported in medial prefrontal cortex⁴ extend to the wider dorsal cortex?* Both questions ask whether prefrontal findings reflect a broader principle of cortical organization, and neither can be addressed with current tools. Tethered miniscopes preclude two unrestrained animals from freely encountering each other in tubes, single-port feeders, or running wheels, and prior wireless devices were too bulky for paired imaging. Simultaneous large-scale dorsal cortical recording from two freely competing mice — what Airscope makes possible — is the experiment that lets these questions be asked.

The added analyses are as follows.

- **Population-level structure of rank.** Projecting trial-aligned cortical activity onto its leading principal components, win and lose trials occupy separable subspaces, with separation detectable before contact (**Fig. 4f-g, Extended Data Fig. 8e-f**). This recapitulates at the dorsal-cortical scale a population-geometry signature previously characterized in mPFC⁴.
- **Interbrain coupling in dorsal cortex.** Cortical activity is correlated above chance across interacting animals. The correlation persists after removing coordinated rest epochs, when only one animal is behaving, and during epochs of low concurrent behavior, ruling out simple co-occurrence of rest or matched movement. This extends the interbrain coupling reported in mPFC (Kingsbury et al., 2019)⁶ to the dorsal cortex.
- **Motor-confound controls in the tube test.** Two complementary controls indicate that rank decoding does not reflect ongoing kinematics. First, removing neurons selective for push, resist, or retreat from the decoder input does not reduce accuracy. Second, fitting a GLM that predicts each neuron's activity from kinematic and behavioral variables and training the decoder on the residual activity also preserves accuracy. Details are in our response to **Major Comment 3, Reviewer 3**.
- **Feeding-confound controls in feeding competition.** The same controls, with eating and port-occupancy variables added to the GLM, preserve decoder accuracy in the single-feeder competition assay.

In parallel, we have added a sensory-region exclusion control to the enriched-environment navigation experiment: removing visual cortical neurons from the decoder input does not impair landscape decoding, indicating that the representation is not carried by visual cortex alone.

Together, these additions change the role of the social interaction experiment in the manuscript. It is no longer a demonstration that a nonlinear decoder can read out the winner from cortical activity; it is a systematic test of two general claims about cortical organization: 1) competitive rank is decodable across dorsal cortex rather than confined to prefrontal circuits, and 2) interbrain coupling is a property of the wider cortex rather than of mPFC specifically. These conclusions are accessible only because both animals can be imaged simultaneously, at scale, while behaving freely.

3. Clearly enabling experiments

We extend the functionality of Airscope by integrating a microLED module for focal optogenetic intervention during simultaneous cortical Ca^{2+} imaging (**Fig. R27a, b**). MicroLEDs have been widely adopted in high-resolution displays, wearable biosensors and, increasingly, in implantable neural interfaces for optogenetic stimulation, owing to their small footprint, high radiant efficiency and addressability at cellular scales^{32,33}. Here, a 100 μm microLED emitting at 570 nm is incorporated to drive yellow-light-activated inhibitory opsins such as Arch or Jaws. The device is powered directly by the Airscope extension PCB through a dedicated power-management circuit that delivers stable current and provides thermal isolation from the imaging electronics, allowing the LED and the image sensor to operate concurrently without optical or electrical crosstalk.

To demonstrate the enabling capability of our integrated platform, we performed simultaneous cortical Ca^{2+} imaging and on-demand optogenetic suppression of epileptic activity as a target application. Epilepsy affects ~50 million people worldwide, and roughly one-third of patients are refractory to pharmacotherapy³⁴. Seizures arise from hypersynchronous neuronal discharges and can progress to status epilepticus—a life-threatening condition requiring immediate intervention. Current therapies, including resective surgery, vagus nerve stimulation and deep brain stimulation, are constrained by incomplete understanding of seizure pathogenesis, lack of cell-type specificity, tissue damage and nonspecific current spread, thus highlighting the need for compact systems enabling real-time detection and focal neuromodulation³⁵. Our wireless head-mounted microscope uniquely meets this need: for the first time, it allows free-moving behavioral assessment essential for scoring seizure severity³⁶ in mouse models, while simultaneously recording large-scale Ca^{2+} dynamics and enabling real-time optogenetic intervention. Such systems hold particular promise for the monitoring and timely mitigation of abrupt, life-threatening episodes, and more broadly for neurorehabilitation in neuropsychiatric disorders.

We validated the integrated imaging-and-intervention system in freely behaving mice ($n = 3$) using the pentylenetetrazol (PTZ) acute seizure model. We observed that cortical seizure waves typically originate from approximately the retrosplenial cortex (RSC) and then spread outward (**Fig. R27c, f**). Notably, because the subiculum (SUB) is an upstream region of the RSC and optogenetic suppression of the SUB has been shown to inhibit epileptic activity^{37, 38}, we selected the SUB as the optogenetic modulation region. For photoinhibition, animals were injected with the virus rAAV-hSyn-ArchT-tdTomato-WPRE-hGHpA at hippocampal SUB. Following intraperitoneal injection of PTZ (2 mg/ml, 0.019 ml/g), seizure progression was scored on the modified Racine scale (**Fig. R27c**): an initial seizure latency (the interval from PTZ injection to the first behavioral manifestation), followed by Racine stage 1–2 events (facial and ear twitching, myoclonic jerks and head nodding), and escalation through stages 3–5 (unilateral forelimb clonus; bilateral forelimb clonus with rearing; and generalized tonic-clonic seizures with loss of postural tone), with mortality in the most severe cases. Seizure latency was defined as the time from PTZ injection to the first Racine level 1 event. In control trials, in which the microLED remained off, animals showed shorter seizure latencies and progressive behavioral escalation on the modified Racine scale (**Fig. R27d, e**). By contrast, optogenetic stimulation significantly prolonged seizure latency (paired t-test, $**p = 0.0068$; **Fig. R27d**). Representative ethograms showed that stimulation delayed the transition from normal activity to Racine level 1 and reduced progression to higher Racine levels compared with paired control trials (**Fig. R27e**). Simultaneous cortical imaging revealed that control trials exhibited a marked increase in frame-wise mean fluorescence intensity during epileptiform activity, whereas optogenetic stimulation maintained cortical activity closer to

baseline during LED-on epochs (**Fig. R27f**). These results demonstrate that the fully wireless Airscope–microLED platform supports simultaneous cortical read-out and focal optogenetic intervention during PTZ-induced seizure, allowing us to directly visualize cortical seizure onset at the RSC, concurrently score seizure severity in freely behaving mice, and confirm that the subiculum (SUB) serves as an effective optogenetic intervention node—a combination of read-out and intervention not achievable with either component alone.

[Figure Redacted]

Fig. R27. Integrated Airscope–microLED platform enables simultaneous cortical Ca^{2+} imaging and optogenetic suppression of PTZ-induced seizures. **a**, Schematic of the Airscope–microLED system for cortical Ca^{2+} imaging and focal 570-nm optogenetic inhibition. **b**, Photographs of the integrated microLED module in the LED-on and LED-off states. **c**, Racine scale used for behavioral seizure scoring, with representative frames of PTZ-induced seizure behavior at level 4. **d**, Relative seizure latency in control and optogenetic-stimulation trials, normalized to the summed latency across paired trials from the same animal (paired t-test, $**p = 0.0068$, $n = 3$ animals). Lines indicate paired individuals; bars show mean $\pm$ SEM. **e**, Representative behavioral ethograms showing Racine-stage progression after PTZ administration in control and optogenetic-stimulation trials. Dashed lines indicate seizure latency (defined as the time from PTZ injection to the first Racine level 1 event). **f**, Simultaneous cortical Ca^{2+} imaging during PTZ-induced epileptiform activity. Top, frame-wise mean fluorescence intensity of cortical recording for control and optogenetic-stimulation groups. Green shaded areas indicate LED-on epochs. Bottom, representative cortical images at the indicated time points (t_1 – t_4). Values on top left denote the corresponding frame-wise mean intensity.

Minor Comment 5 — Awkward sentence constructions

There are many weird sentence constructions that do not make sense such as: 444 Here, to our knowledge, we used the Airscope system to measure cortical activity during 445 large-scale labyrinth navigation (Fig. 5a and Supplementary Fig. 19a-c)

Response: We sincerely apologize for the unclear language. We have carefully proofread the entire manuscript and corrected all grammatically awkward constructions. The specific example has been revised to: "*We used Airscope to measure cortical activity during large-scale labyrinth navigation.*" The manuscript has been additionally reviewed by a native English-speaking colleague.

Minor Comment 6 — Faster imaging rate demonstration

If possible, to show a system with a faster imaging rate would significantly enhance the capability of the system, even at the cost of some additional weight or in a wired mode in case the wireless transmission is the limiting step for faster imaging.

Response: We thank the reviewer for this suggestion. In the revised manuscript we have implemented and characterized a higher-speed configuration of Airscope.

By applying on-sensor binning on the OV2640, we raised the frame rate to 30 Hz at 800×600 pixels (**Fig. R28f-h**), matching the temporal resolution of Kiloscope noted by the reviewer. Binning necessarily reduces the spatial sampling density; within the revision timeline it was not feasible to fabricate a new higher-magnification lens set to recover the lost sampling resolution. Instead, we inverted the optical layout of the existing Airscope to increase the effective magnification and achieved a lateral resolution of $\sim 2.19 \mu\text{m}$, corresponding to group 7, element 6 of the USAF-1951 target (**Fig. R28e**). Critically, the acquisition board, firmware and software are unchanged from the original Airscope, so this higher-resolution configuration is a drop-in modification of the same platform rather than a redesign.

We respectfully note that wireless operation is a core design principle of Airscope; introducing a wired data link, even transiently to secure a higher frame rate, would negate the primary advantage that distinguishes this platform from existing miniaturized microscopes. The configuration we report demonstrates that higher temporal resolution is achievable within the wireless framework, without compromising the untethered paradigm that motivates the work.

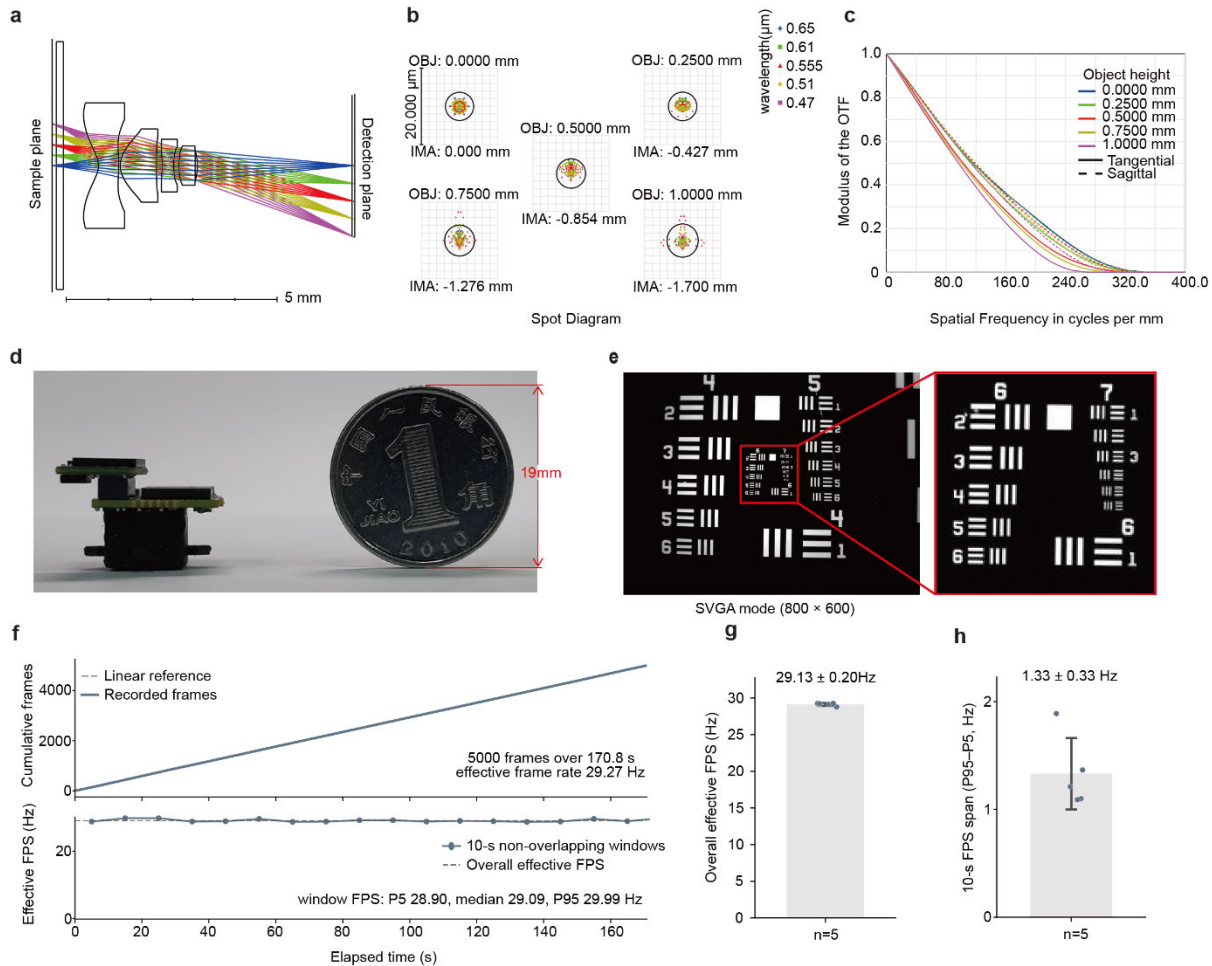


Fig. R28. High-frame-rate imaging with a reversed high-magnification camera-module configuration. **a**, Optical layout of the reversed OV2640-based configuration, which increases magnification from $\sim 0.5\times$ to $\sim 2\times$. **b**, Spot diagrams at representative field positions. **c**, Image-space MTF at corresponding object heights. **d**, Photograph of the assembled module next to a 19-mm-diameter coin for scale. **e**, USAF 1951 resolution target image acquired in native SVGA mode (800×600), with the inset showing resolvable features up to Group 7, Element 6. **f**, Representative timestamp-based continuity analysis showing cumulative frame acquisition and effective FPS in 10-s non-overlapping windows. **g**, Across-device summary of overall effective FPS. Bars indicate mean and error bars indicate SD; dots denote individual devices. **h**, Across-device summary of 10-s window FPS spread, defined for each recording/device as the P95–P5 span of effective FPS values from 10-s non-overlapping windows. Symbol as in **g**.

Minor Comment 7 — Zemax file release for benchmarking

Providing Zemax files would give this work the possibility to become a reference and benchmark for all future designs and would certainly maximize impact

Response: We sincerely thank the reviewer for highlighting this opportunity. We have now released the complete Zemax design files (.zmx), including lens prescriptions and optimization merit functions, as part of our comprehensive open-source package. We hope this release will serve

as a useful reference for the community and facilitate future innovations in miniaturized microscope design.

Reviewer #3

Remarks to the Author

In this manuscript, a new miniature wireless one-photon microscope with large field of view is developed and used in mice as they engage in various behaviors where imaging with a tethered miniature microscope or a tabletop microscope would be difficult or impossible. In addition to the basic introduction of the new technology, the utility and performance of this new device is showcased during exploration in a naturalistic habitat and various other setups. Together with the hardware, multiple software tools are developed, including for neuron segmentation, behavior tracking, decoding activity, and other purposes. The hardware development on the side of optics, onboard processing, wireless transmission, and integration across systems is impressive.

Overall, this is a very impressive suite of work, that likely will be highly impactful for systems neuroscience, including for recording neural activity while mice explore in outside spaces.

We thank the reviewer for this enthusiastic and encouraging assessment. We deeply appreciate the recognition of the engineering effort invested in optics, onboard processing, wireless transmission, and system integration. We have worked to address all comments comprehensively in our revision.

In addition to this enthusiastic overall assessment, below are a number of comments and concerns:

We thank the reviewer for the thoughtful and constructive comments that follow. Each point has been carefully addressed, and we believe the manuscript is substantially strengthened as a result.

Major comments:

Major Comment 1 — Accessibility of technology description for non-expert readers

In my opinion, the text and figures somewhat underemphasize the strength of the technology development and could provide readers new to the field with more intuition about how this progress was possible. For example, on the side of optics, it is described only briefly why the Airscope optical design allows for large field of view imaging using a short optical path length, and explains little about the relationship to the Chevalier landscape lens, how it was manufactured, and so on. These sections do not need to be longer, but they could be made more accessible to non-expert readers.

Response: We thank the reviewer for this constructive suggestion. We have made three coordinated revisions, designed to deepen the physical intuition without lengthening the Results section.

(i) Rewritten optics paragraph in Results. The paragraph now explicitly conveys the three intuitions the reviewer requested. The revised text reads:

"The Airscope optical module integrates four aspheric elements that deliver 6-mm FOV at 4-6 μ m resolution within an exceptionally short image-plane-to-sensor distance (Fig. 1c and Supplementary Fig. 1). Following the classic Chevalier landscape architecture, the aperture stop is placed in front of the imaging stack: this front-stop asymmetry routes off-axis chief rays so that residual aberrations partially self-cancel inside the glass, performing optical work that would otherwise require additional elements. Aspheric surfaces add the surface-shape degrees of freedom needed to jointly correct coma, spherical aberration, astigmatism, and Petzval field curvature within only four elements, while halving the conjugate distance relative to all-spherical alternatives (Supplementary Fig. 2). The full prescription was generated by a custom curriculum-

learning optical-design agent (Supplementary Note 1), and the elements are mass-produced by injection molding at <\$2 per module (Supplementary Table 3). The resulting 0.1-g, $7 \times 7 \times 3 \text{ mm}^3$ pack achieves a space-bandwidth product among the highest reported for head-mounted optics (Supplementary Table 1 and Methods). The depth of field of the Airscope is established at $\sim 90 \mu\text{m}$, providing sufficient axial coverage to encompass the majority of cortical layer 2/3 within the target animal model (Methods, Supplementary Fig. 8)."

Specifically, the new wording (a) explains *why* aspheric surfaces correct field curvature and off-axis aberrations — they supply the additional surface-shape degrees of freedom needed to balance four aberration types within only four elements; (b) makes the Chevalier relationship explicit — the front-stop asymmetry is *borrowed* and provides aberration self-cancellation, while the four aspheres and ultrashort conjugate are our *modifications*; and (c) names the manufacturing route — high-precision injection molding, which enables mass production at < \$2 per module.

(ii) New Supplementary Note 1. For readers who wish to go further, **Supplementary Note 1** describes the full LensCopilot curriculum-learning design framework that generated the prescription, including staged release of aspheric degrees of freedom, manufacturability constraints, and the Zemax local–Hammer–local optimisation loop. The same Note also serves **Reviewer 2's Major Comment 1**; we cross-reference rather than duplicate.

Together, these three changes address the reviewer's request for accessible intuition while preserving the conciseness of the main text.

Major Comment 2 — Move microscope characterization (axial resolution) to main text

The demonstrations of the utility of the technology in the main text are extensive, but the characterization of the microscope itself is not as extensive and is covered in more detail in the Supplementary Material. It would be useful to move some of that characterization to the main text. Especially the axial resolution, measured in Supplementary Figure 10b, is useful to include, because it shows that signals from neurons at different depths can overlap and have to be recovered using algorithms like DeepWonder, used in the manuscript.

The larger FWHM axially compared to laterally is expected as this is a 1-photon system, and to my knowledge it is often not shown in 1-photon miniscope publications, so it is good that the authors characterized it here.

Response: We thank the reviewer for this constructive suggestion and for the positive acknowledgment regarding the axial characterization, which is indeed rarely reported in 1P miniscope publications.

We have made the following revisions:

1. Axial resolution in the main text. The axial point-spread function measurement, or the depth of field, (previously **Supplementary Fig. 10b**) has been promoted to a main-text figure panel (revised **Fig. 1c**, see also in the **Fig. R29**). This addition provides a more complete characterization of the optical performance of Airscope and now explicitly illustrates the broader depth of field inherent to 1P excitation. This extended depth of field enables Airscope to record from a larger axial volume and thereby capture more neurons within the imaged region, it also introduces axial crosstalk because neurons at different depths can be projected onto overlapping locations on the 2D sensor. So in the following manuscript text, we explicitly state that the DeepWonder

framework was employed to resolve neurons that are physically separated along the z-axis but become superimposed on the 2D image sensor due to the extended axial point spread function (PSF): “A computational pipeline extending the DeepWonder framework suppresses neuronal-neuropil crosstalk, rejects out-of-focus fluorescence, and segments laterally overlapping neurons across the depth of field (Fig. 2f and Extended Data Fig. 4)” (Section 2 of the revised manuscript). The signal demixing capability of the DeepWonder framework is further demonstrated in this document (see **Fig. R16** and the response to **Reviewer 2, Major Comment 3**). In that experiment, we evaluated the algorithm's ability to resolve laterally overlapping sources, demonstrating that temporal correlation scores remain robust even when the intercellular distance between two neurons is reduced to 0.3 somatic diameters.

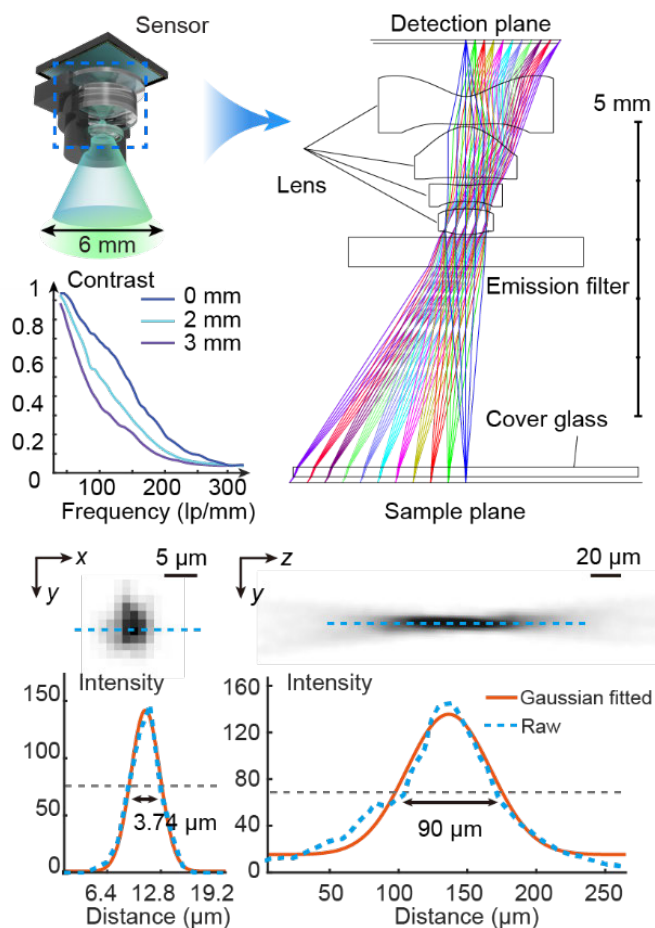


Fig. R29. Revised Fig. 1c to provide a more complete characterization of the optical performance of Airscope. Top (original figure): schematic of the Airscope optical module and ray-tracing simulation from the sample plane to the detection plane and modulation transfer function (MTF) at different lateral displacements from the FOV center. Bottom (additional part): representative lateral and axial intensity profiles of a 1 μm point emitter captured using the Airscope lens coupled to a high-resolution 200 MP sensor.

2. Two-photon ground-truth validation of signal overlap. To further validate the effectiveness of the DeepWonder-based pipeline in mitigating crosstalk caused by overlapping neuronal signals, we coupled a two-photon microscope to the Airscope system and performed temporally

interleaved Airscope imaging under one-photon illumination and two-photon imaging in the same regions. This hybrid configuration provided near-simultaneous two-photon ground truth at selected depths beneath the dura, allowing us to evaluate both spatial localization and temporal activity recovery from the Airscope data. The paired recordings showed high consistency in neuronal locations and activity traces across the two modalities (**Fig. R17b, c, d**), with a temporal correlation of 0.82 ± 0.14 across 1,196 neurons (mean $\pm$ SD). Additional analyses across multiple cortical regions showed consistent correlation scores (**Fig. R18**), further supporting robust neuronal signal demixing by the DeepWonder-based pipeline.

We also used this paired 2P–Airscope system to assess depth-dependent performance. Recordings were acquired at multiple depths beneath the cortical surface, including 100, 150, and 200 μm (**Fig. R20**). Although temporal correlation decreased modestly with increasing imaging depth, likely due to increased scattering and axial signal integration, Airscope maintained a correlation of 0.82 ± 0.15 at 200 μm depth (**n = 102 neurons**, mean $\pm$ SD). Neuronal detection accuracy across depth, quantified by the F1 score, was 0.75 ± 0.02 across three animals (mean $\pm$ SD).

This validation confirms that our processing pipeline successfully recovers individual neuronal signals from the overlapping data. These results are presented in the Extended Data Fig. 5 and more details in the response to the Reviewer 2.

Major Comment 3 — Over-interpretation of biological results

A recurrent (but solvable) issue with the manuscript is the slight **but pervasive over-interpretation of the biological results**, as occurs for almost all experiments.

Response: We sincerely thank the reviewer for this careful and constructive feedback. We fully agree that precise language is essential, particularly in a methods-focused manuscript where the biological demonstrations serve primarily to showcase the technology's capability rather than to establish definitive neuronal mechanisms. We have conducted a thorough revision of the entire manuscript to ensure all claims are commensurate with the supporting evidence. Below we address each specific example:

3a — Social hierarchy decoding: motor confounds

For instance, in the decoding of social hierarchy, the manuscript claims that there is a cortex-wide representation of hierarchy, and posits that a previous result suggesting this was encoded in prefrontal cortex is incomplete. While it is plausible that indeed these representations are cortex-wide, the evidence presented does not show this. That is because the neural activity may contain correlates of the different types of motor actions performed by the animals, which are different for the dominant and subordinate mice. For example, the dominant mouse probably pushes more and the subordinate mouse walks backward more, allowing for purely motor representations to be informative about social rank, even if they do not explicitly encode rank. To claim that the cortical representations encode place in the social hierarchy would require very careful disambiguation from differences in motor actions and sensory experiences. Therefore, this claim should be toned down.

Response: We thank the reviewer for this insightful critique. We fully agree that the original language overstated what the decoding results demonstrate, and we acknowledge that motor and sensory confounds must be explicitly discussed.

In the revised manuscript, we have:

1. Revised all claims to observational language. "Cortex-wide representation of social hierarchy" has been replaced with: *"Cortical activity patterns across multiple dorsal regions correlate with behavioral signatures associated with social dominance interactions."* in the abstract. We now explicitly frame the decoding results as demonstrating that Airscope captures distributed cortical signals that covary with social behavioral states, without claiming these signals encode rank per se.

2. Evaluated the contribution of motor-related activity to decoding.

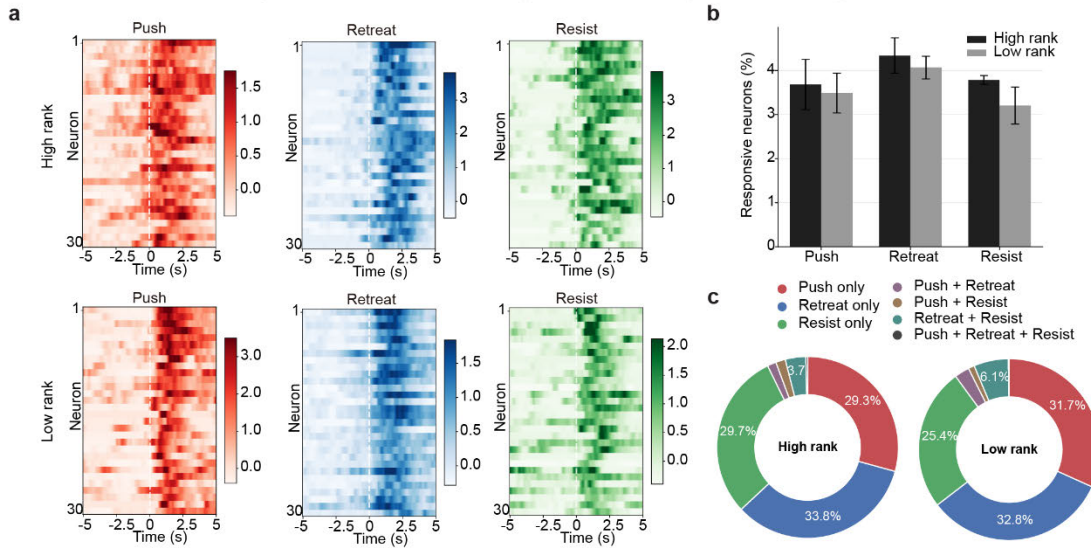
We added a dedicated motor-confound analysis to examine whether win–lose decoding could be explained by overt behavioral differences during tube-test interactions. We first quantified neurons responsive to the main motor behaviors observed during the task, including push, retreat and resist. As shown in **Fig. R30a-c**, subsets of neurons were significantly responsive to these behaviors in both high-rank and low-rank animals. We then performed two complementary analyses to assess whether such motor-related activity accounted for the decoding performance.

First, neurons significantly tuned to push, retreat or resist were excluded before training win–lose decoders (**Fig. R30d**). Removing these behavior-tuned neurons did not significantly affect decoding performance (**Fig. R30e**; unpaired t-test; $p = 0.4257, 0.3696$ and 0.2675 for accuracy; $p = 0.7534, 0.0986$ and 0.4481 for F1 scores, for SVM, MLP and Dual-stream BERT classifiers, respectively). This result suggests that win–lose decoding was not driven solely by neurons individually tuned to these overt motor behaviors.

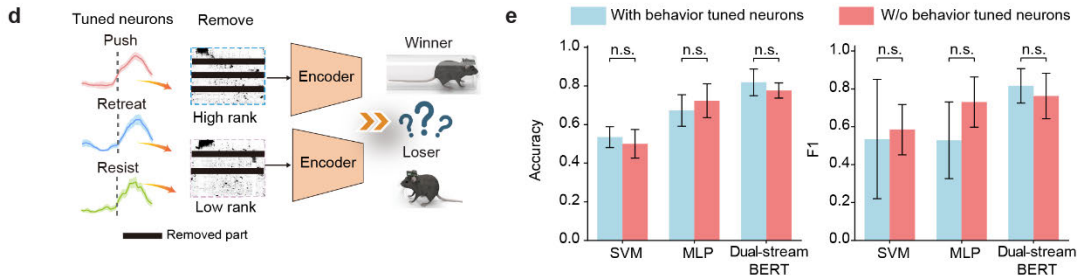
Second, to further assess the contribution of population-level motor-related signals, behavioral variables including still, push, retreat and resist are used to fit neuronal population PCs with Generalized Linear Model (GLM) (**Fig. R30f**). The GLM explains a modest fraction of PC variance in both high-rank and low-rank animals (**Fig. R30g-h**). GLM-fit residuals are calculated as the difference between the original neuronal PC activity and the GLM-fitted activity (**Fig. R30i**), and win–lose decoding is repeated using these residuals. Decoding performance is not significantly reduced after removing GLM-fitted motor-related components (**Fig. R30j**; two-sided independent samples t-test; $p = 0.9182, 0.7478$ and 0.3181 for accuracy; $p = 0.7265, 0.2466$ and 0.5396 for F1 scores, for SVM, MLP and two-stream classifiers, respectively). These results suggest that the win–lose decoding performance was not trivially explained by the measured motor variables.

These analyses have been added to the revised manuscript in **Extended Data Fig. 8**.

Specific behaviors in social competition evokes single-neuron tuning



Excluding motor confound neurons does not affect win/lose decoding



Excluding motor confound population factors does not affect win/lose decoding

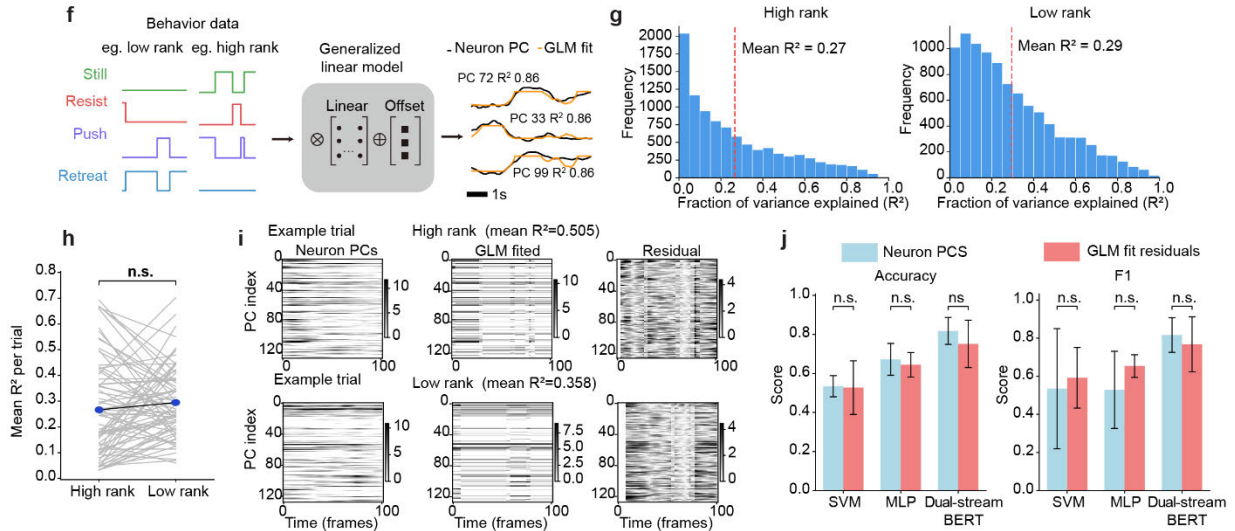


Fig. R30. Motor confounds in social hierarchy decoding. **a**, Trial averaged responses of example neurons tuned to push (red), retreat (blue), and resist (green) for high rank and low rank animals. **b**, Percentage of neurons significantly responsive to each behavior for high-rank and low-rank animals. Bar plots indicate mean $\pm$ s.e.m. **c**, Proportions of neurons selectively or jointly responsive to push, retreat and resist behaviors in high-rank and low-rank animals. **d**, Schematic of the motor-confound control analysis through excluding motor-tuned neurons. Neurons tuned to push, retreat

or resist behaviors are removed before training decoders to classify win–lose outcome. **e**, Decoding accuracy and F1 scores for SVM, MLP and Dual-stream BERT classifiers with or without behavior tuned neurons (unpaired t-test; $p = 0.4257, 0.3696$ and 0.2675 for accuracy; $p = 0.7534, 0.0986$ and 0.4481 for F1 scores). Bar plots indicate mean $\pm$ SD across five-fold validation. **f**, Schematic of the population-level motor-confound regression analysis. Behavioral variables including still, push, retreat and resist are used to fit neuronal population PCs via GLM. Motor-related components are estimated from the GLM-fitted activity and removed to obtain GLM-fit residuals for subsequent win–lose decoding. **g**, Distribution of the fraction of variance explained by behavioral variables for neuronal PCs in high-rank and low-rank animals. Dashed red lines indicate the mean explained variance across PCs (high rank: mean $R^2 = 0.27$; low rank: mean $R^2 = 0.29$). **h**, Comparison of the mean behavioral variance explained per trial between high-rank and low-rank animals. Grey lines indicate paired trials and blue dots indicate group means. No significant difference was observed between ranks (two-sided Wilcoxon rank-sum test, $p = > 0.1$). **i**, Example trial showing neuronal PCs, GLM-fitted activity and GLM-fit residuals for high-rank and low-rank animals. GLM-fit residuals are calculated as the difference between the original neuronal PC activity and the GLM-fitted activity. **j**, Decoding accuracy and F1 scores for SVM, MLP and two-stream classifiers using original neuronal PCs or GLM-fit residuals (two-sided independent samples t-test; $p = 0.9182, 0.7478$ and 0.3181 for accuracy; $p = 0.7265, 0.2466$ and 0.5396 for F1 scores). Bar plots indicate mean $\pm$ SD across five-fold validation.

3b — Social feeding setup: taste and eating confounds

A similar comment applies to the experiment with the social feeding setup. Here, the cortical activity patterns that allow social status to be decoded may be neurons representing eating or taste, for example. Similarly, the claims are overstated as they are written now, and should be toned down

Response: We thank the reviewer for this thoughtful comment. We agree that, in the social feeding assay, cortical activity patterns used to decode social status could be influenced by feeding-related variables. We have therefore revised the interpretation of this experiment and added new analyses to evaluate the contribution of feeding-related confounds.

In the revised manuscript, we have:

- 1. Revised the language to avoid overstatement.** We have applied the same systematic language revision to the social feeding experiment as in question 3a.
- 2. Evaluated the contribution of feeding-related behavioral variables.** To directly assess whether feeding-related behavior could account for the observed neural activity patterns, we extract behavioral variables related to food-port interaction, including occupancy ratio, distance to the feeder port and occupancy time (**Fig. R31a**). We then fit neuronal population PCs using these feeding-related behavioral variables via GLM (**Fig. R31b**). The feeding-related variables explain only a limited fraction of neural population variance across components in both high-rank and low-rank animals (**Fig. R31c**), and the mean explained variance per session is not significantly different between ranks (**Fig. R31d**). The GLM reconstructed rastermap of neuronal activity PCs failed to recapitulate the structure of empirical PCs (**Fig. R31e**).
- 3. Tested whether feeding-responsive neurons dominate rank decoding.** We next identify neurons that are strongly correlated with feeding episodes and show repeated activity increases

during eating events (**Fig. R31f**). Feeding-responsive neurons are present in both high-rank and low-rank animals (**Fig. R31g**). To test whether these neurons drove decoding performance, we remove feeding-tuned neurons and re-evaluate feeder occupancy prediction without changing the trained classifier framework. Decoding accuracy and F1 scores changed only minimally after excluding these neurons across SVM, MLP and Dual-stream BERT classifiers (**Fig. R31h**). These results suggest that the decoding performance was not dominated by feeding-tuned neurons alone.

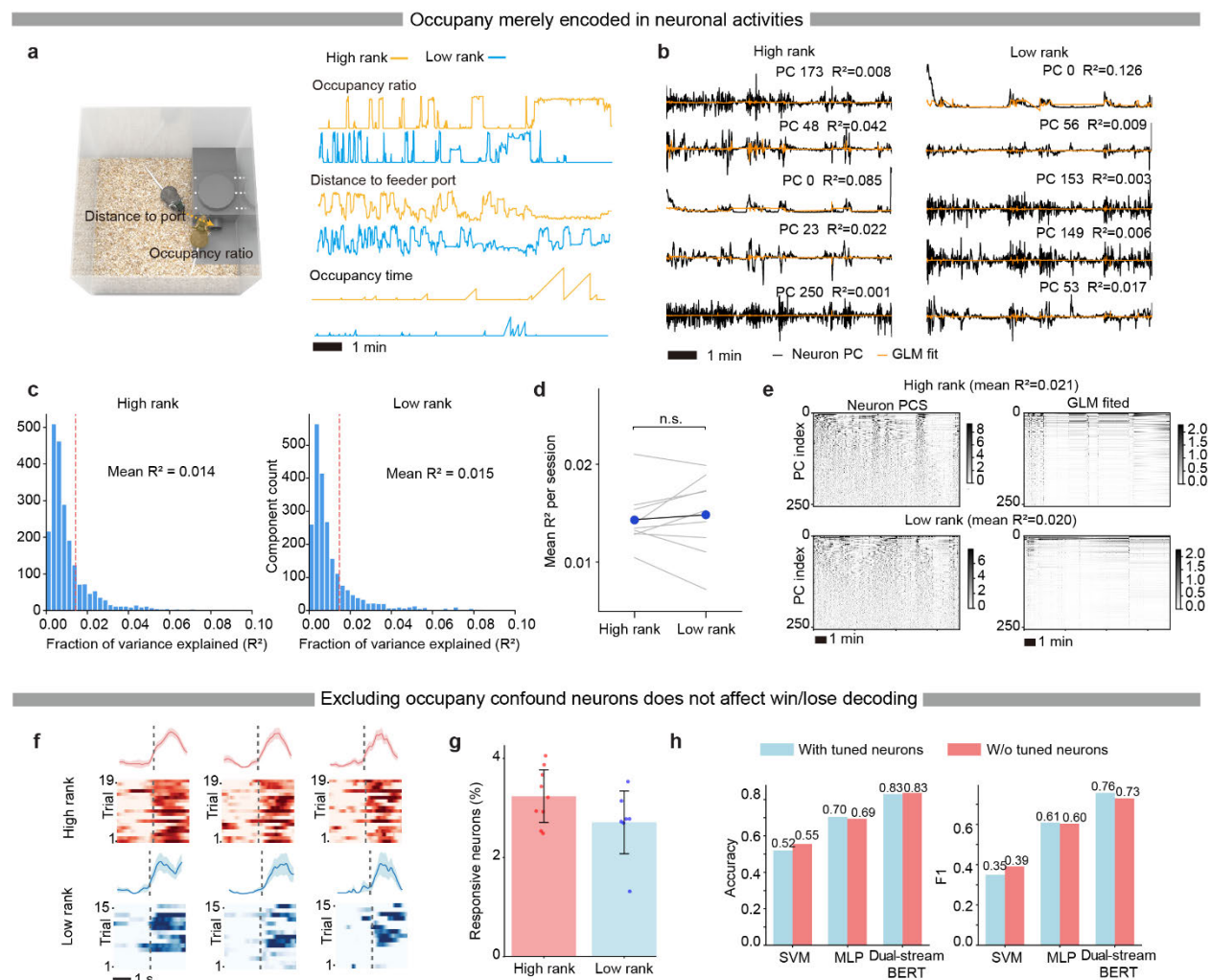


Fig. R31. Feeding-related activity and decoding performance in the social feeding assay. **a**, Schematic of the feeding competition assay (left) and example traces of feeding-related behavioral variables from high-rank and low-rank animals (right). **b**, Random selected neuronal PCs fitted by GLM using feeding-related behavioral variables for high-rank and low-rank animals. Black traces indicate neuronal PC activity and orange traces indicate GLM-fitted activity. **c**, Distribution of the fraction of variance explained by feeding-related behavioral variables for neuronal PCs in high-rank and low-rank animals. Dashed red lines indicate the mean explained variance across components. **d**, Comparison of the mean feeding-related variance explained per session between high-rank and low-rank animals. Grey lines indicate paired sessions and blue dots indicate group means. No significant difference was observed between ranks (two-sided Wilcoxon rank-sum test, $p = 0.66$). **e**, Example session showing neuronal PCs and GLM-fitted PCs for high-rank and low-

rank animals. **f**, Representative event-aligned neuronal activity during feeding episodes in high-rank and low-rank animals. Heatmaps: trial-wise responses; center line: mean; shaded area: s.e.m. **g**, Percentage of feeding-responsive neurons in high-rank and low-rank animals. Bars show mean $\pm$ SD; dots indicate individual animals. **h**, Decoding accuracy and F1 scores for SVM, MLP and Dual-stream BERT classifiers with or without feeding-tuned neurons.

3c — Landscape exploration decoding: low-level visual confounds

For the landscape exploration decoding, where different zones in the landscape can be decoded from cortical activity, the same comment applies. Whether **these are encoding 'landscape-specific cortical representations' in a more abstract way, or something lower level like average color, different spatial (visual) frequencies, or possibly location, cannot be deduced without careful control experiments**. Again, the claims should be changed to what can actually be concluded. There are other cases throughout the manuscript where the biological claims should be revisited and possibly reworded in a more concrete way.

Response: We thank the reviewer for this consistent and constructive critique. We have revised the landscape exploration section accordingly:

1. Revised claims. "Landscape-specific cortical representations" has been removed. We carefully removed all descriptions suggesting that zone identity decoded from cross-zone cortical neuronal activities reflects abstract landscape representations.

2. Evaluated the contribution of visual cortical activity to landscape-zone decoding. To assess whether landscape-zone decoding was dominated by activity from visual cortical neurons, we quantify the proportion of recorded neurons from each visual cortical area, including VISa, VISam, VISpm, VISp and VISrl (**Fig. R32a**). We then repeat landscape-zone decoding after excluding visual cortex neurons and compared the decoding performance with the original model using all recorded neurons. Removing visual cortex neurons does not significantly reduce decoding performance (**Fig. R32b**), which suggests that the landscape-zone decoding was not driven solely by neurons recorded from visual cortical areas. We have reported this control in the main manuscript as "*Excluding neurons in visual cortex did not degrade decoding performance (Fig. 3h), arguing against a purely visual-feature account.*" The control figure has been integrated into **Fig. 3h** in the revised manuscript.

3. Full-manuscript language review. In addition to the three specific examples addressed above, we have conducted a comprehensive review of the entire manuscript and revised all biological claims to ensure they describe observed correlations and demonstrated decoding capabilities, rather than implying mechanistic understanding.

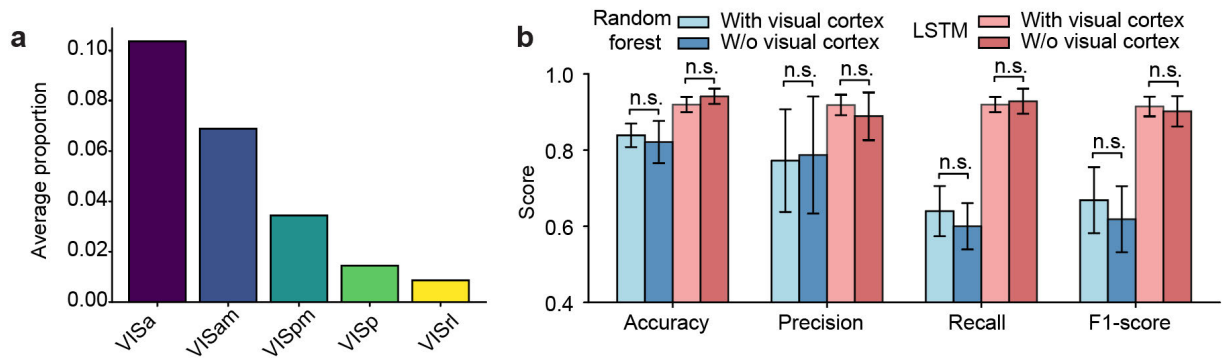


Fig. R32. Contribution of visual cortical activity to landscape-zone decoding. **a**, Average proportion of neurons recorded from each visual cortical area ($n = 3$ animals). **b**, Decoding performance with or without visual cortex neurons (two-sided Wilcoxon rank-sum test, $n = 3$ animals). Random forest: accuracy, $p = 0.70$; precision, $p = 1.00$; recall, $p = 0.70$; F1-score, $p = 0.70$. LSTM: accuracy, $p = 0.40$; precision, $p = 0.70$; recall, $p = 1.00$; F1-score, $p = 1.00$). Bars show mean $\pm$ SD; n.s., not significant.

Positive acknowledgment from the reviewer

Importantly, these above comments about over-interpretation of the data are not negative, because nobody would expect a single methods-focused study to discover all these cortical representations. It is already impressive that these task and behavioral variables can be decoded from the data using a single field of view, and an adequate demonstration of the utility of the technology

We sincerely appreciate the reviewer's balanced perspective and encouraging framing. We agree that the primary contribution is technological, and we have restructured the narrative to position all biological experiments as demonstrations of Airscope's versatility and data quality, while reserving mechanistic interpretation for future dedicated studies.

Major Comment 4 — Layer specificity and depth adjustment mechanism

Layer specificity: the microscope is designed to record from a fixed depth and has, to my understanding, no option to change the imaging plane after mounting. This should be clearly stated. It does raise the question of whether there is a mechanism to adjust the vertical position before fixing the microscope, or whether the setup can be extended with an integrated simple vertical displacement such as a miniature stage or screws to adjust the height afterwards. It would be useful to at least discuss this

Response: We thank the reviewer for raising this important practical concern, and we have addressed it on two complementary fronts.

1. Pre-mounting depth adjustment via spacers (existing mechanism). We would like to clarify that the current design does provide a mechanism to set the vertical position before the microscope is fixed. Specifically, we have designed a set of spacers with 20 μm height increments that are inserted between the baseplate and the Airscope body. Selecting different spacers shifts the Airscope vertically, thereby tuning the imaging depth before microscope fixation. In practice, the appropriate spacer is determined once per animal to reach the optimal imaging depth, and is then kept fixed so that the same plane can be recorded reproducibly across sessions. The spacer design

files have been released together with the other Airscope design files in our open-source repository. In the revised manuscript, we have added a transparent statement of this limitation as part of Discussion, which now explicitly reads: *"In the standard configuration of Airscope, the imaging plane is set by the chosen spacer and cannot be adjusted after the microscope is mounted; further mechanical variants can address this constraint."*

2. New post-mounting depth-tuning module. Motivated by the reviewer's suggestion, we have leveraged the revision period to develop and validate an integrated depth-tuning module (**Fig. R33a**). The mechanism consists of a miniature precision screw system incorporated into the baseplate design, enabling $\pm 400\ \mu\text{m}$ vertical displacement of the focal plane after implantation (**Fig. R33b**). Importantly, the module provides continuous depth tuning without discrete step-size limitations, allowing the experimenter to freely adjust the imaging plane to target different cortical layers across sessions. The module adds only 0.4 g to the total system weight and allows the experimenter to adjust the imaging depth to target different cortical layers across sessions. This new capability directly addresses the reviewer's concern and enhances the versatility of the Airscope platform for multi-layer cortical studies.

[Figure Redacted]

Fig. R33. Depth tuning of Airscope. **a**, Schematic of the integrated depth-tuning module before mounting, showing the focus ring, retaining ring, baseplate and locking screw. Rotation of the focus ring produces axial displacement of the optical assembly, providing a $\pm 400\ \mu\text{m}$ focal-plane tuning range after implantation. **b**, Validation of post-mounting depth tuning. Images of the mounted module and corresponding resolution-target images are shown at three focal-plane positions spanning -400, 0 and 400 μm .

Major Comment 5 — Integrated software package for analysis

That the manuscript offers not only a novel microscope but also a set of developed algorithms that can be used to get the most benefit from the microscope can be described more clearly. For example, the authors could create an integrated package that brings the behavioral and image analysis together in one place for increased adoption of the entire suite of methods

Response: We thank the reviewer for this valuable suggestion, which highlights the importance of software accessibility for community adoption.

1. Improved software description. We have revised the manuscript to more clearly describe the full suite of developed algorithms — including calcium imaging processing pipeline, SAM2Mice, and the BERT like predictor — as an integrated analytical framework designed specifically for Airscope data.

2. Unified repository and documentation. We have created a centralized GitHub repository (Airscope_toolbox) that serves as the single-entry point for all software tools, with cross-linked documentation, example datasets, and step-by-step tutorials. Each tool is available as a standalone module, and we provide example Jupyter notebooks demonstrating the full end-to-end pipeline from raw Airscope data to behavioral and neural analysis outputs.

Major Comment 6 — Figure 1a clarity and panel 1m background

Some comments on the figures for clarity. Figure 1a is confusing because it is unclear what the arrows mean. Readers might assume, for example, that the first panel represents the implanting of the Airscope, the middle panel a wired mode, and the third panel the wireless mode. It should be clarified that the first two panels are not what the manuscript is about. Regarding panel 1m, the background appears to be replaced by a color gradient, which at first made me think this was an illustration instead of a real photograph. Can the legend clarify that the background was replaced, or can the full photograph be shown.

Response: We thank the reviewer for these specific and helpful suggestions regarding figure clarity.

1. Figure 1a redesign. We have redesigned **Fig. 1a** to eliminate the ambiguous arrow-based layout. The revised panel now uses a clear two-part structure: the left side shows prior approaches (tethered tabletop microscope, wired miniscope) with a visual boundary, and the right side introduces Airscope as the wireless alternative. Explicit labels and a contrasting color scheme ensure that readers immediately understand the first two panels represent existing technologies, not the Airscope system.

2. Panel 1m photograph. We have replaced the gradient-processed background with the original unedited photograph (now in panel 1l). The figure legend now states: "*Photographs of mice engaged in classical behavioral paradigms within ecological environments while their brain activity was continuously recorded. Paradigms shown include rope climbing and walking through a narrow tube. No image processing was applied to the photograph.*"

Major Comment 7 — Figure 2b/2c: choice of showcase images

In Figure 2b and 2c, the first two images showcasing the quality of the Airscope show blood vessels and microglia. This seems like an odd choice, given that the primary use of the Airscope is to image neural activity. A raw image of neurons is shown in 2f, but it is smaller and later. It seems that for a paper focused on neurons, neuron tissue should be shown first and in full detail.

Response: We thank the reviewer for this suggestion regarding figure presentation. We would like to clarify that original Figure 2b actually presents Thy1-GFP labeled neuronal fibers, not blood vessels. The purpose of 2b and 2c is to show comparable resolution between Airscope and benchtop macroscope in neuronal imaging.

Nevertheless, we agree that the original panel arrangement did not effectively communicate Airscope's primary function. Thereby, in the revised Figure 2, we have reorganized the panel as suggested by the reviewer.

1. Original 2c and 2e were moved to the **Supplementary Fig 9**, which were replaced by neuronal fluorescence images shown at full display size, in panel c. GCaMP6f-labeled cortical neurons with

clearly visible individual somata, with comparison with a benchtop microscope. Statistically comparison of calcium event rate between Airscope and benchtop microscope is shown in panel e.

2. Original Figure 2b is maintained, as we thought it is a good case to demonstrate the resolvability in vivo resolving the fine structure, demonstrating the in vivo resolution.

This enhancement immediately conveys to the reader that Airscope is primarily a functional neural imaging tool.

Major Comment 8 — GitHub repository not accessible

The GitHub link, https://github.com/Neurallabware/Airscope_toolbox, appears not to work, so this code could not be reviewed. Perhaps the repository has not been set to public yet

Response: We sincerely apologize for this oversight. The repository was inadvertently left in private mode during the initial submission. It has now been set to public and is fully accessible at: https://github.com/Neurallabware/Airscope_toolbox.

The repository now includes the complete Airscope codebase, including firmware and acquisition-related resources, calcium-imaging analysis pipeline, SAM2Mice for multi-animal segmentation, the BERT-based social competition predictor, and pretrained model checkpoints. Installation instructions, dependency lists, example datasets, and Jupyter notebooks for loading, visualization, processing, and benchmarking are also provided.

In addition, we have expanded the accompanying online documentation resources:

- <https://airscope.org/devkit/> — hardware assembly, firmware flashing, acquisition setup, and platform-extension instructions.
- <https://airscope-docs.readthedocs.io/en/latest/> — complete documentation for Airscope hardware, software, datasets, and analysis workflows, including usage examples for the calcium-imaging pipeline, SAM2Mice, and the BERT-based predictor.

We note that both Referee #1 and Referee #2 independently flagged this issue, confirming the importance of code availability. We have verified that all links in the manuscript and supplementary materials resolve correctly.

Major Comment 9 — Open-source material listing clarity

The optics, PCB board designs, firmware, and so on are available as open source as listed under the Data availability section. This is great but it is not immediately clear because only "CAD models" is mentioned. It would be beneficial to showcase more clearly what is in the google drive. Can you list the items so readers know that the full design is there

We thank the reviewer for this practical suggestion. We have revised the Data Availability section to provide a comprehensive, itemized list of all shared materials:

- **Optical design:** Zemax files (.zmx) with full lens prescriptions and optimization merit functions.
- **Electronics:** PCB schematics and layout files, bill of materials.

- **Mechanical:** CAD models for all housing, baseplate, and depth-tuning components (STEP and STL formats).
- **Firmware:** Source code for onboard MCU and wireless transmission protocol.
- **Software:** Full analysis pipeline including calcium-imaging analysis pipeline, SAM2Mice, and BERT-based predictor.
- **Documentation resources:** Assembly instructions, alignment procedures, firmware flashing procedures, acquisition-software setup, and platform-extension documentation is available at <https://airscope.org/devkit/>. Full online documentation for the Airscope hardware, software, datasets, and analysis workflows is available at <https://airscope-docs.readthedocs.io/en/latest/>.
- **Data resources:** Including raw calcium-imaging recordings, multi-animal interaction and enriched-habitat datasets, SAM2Mice benchmark data, optical calibration data (see https://github.com/Neurallabware/Airscope_toolbox/tree/main/Data_release/optical_characteristic), and MCU timestamp data, together with example Jupyter notebooks for data loading, visualization, processing, and benchmarking.

This itemized list is now included in the Data Availability section of the manuscript, and a README file in the Google Drive root directory (https://drive.google.com/file/d/1CB_UG3Mj7FI_y-4qctgCWbV-mwNIDweI/view?usp=drive_link) provides an overview of the release structure.

Minor comments

Minor Comment 1 — Abstract language: "track length" and "tabletop-grade throughput"

The abstract contains some language that is hard to understand. "track length" (line 35) should be clarified to mean optical path length from the sensor to the edge. "tabletop-grade throughput" should be clarified to mean the number of neurons recorded using a tabletop one-photon microscope.

Response: We thank the reviewer for identifying this unclear language. We wish to clarify that track length—or more precisely, total track length (TTL)—is a standard optical engineering term denoting the physical distance from the anterior surface of the primary lens element to the final image plane (sensor). It is a critical determinant of overall system compactness. To prevent future confusion, we have added a dedicated subsection to the Methods defining these spatial parameters:

“Optical path metrics. Total track length (TTL) and conjugate length denote distinct spatial parameters within the optical design. TTL is the physical distance from the anterior surface of the primary lens element to the final image plane (sensor). As it encompasses both the physical footprint of the lens assembly and the back focal length, TTL serves as a critical determinant of overall system compactness. Conversely, conjugate length defines the total spatial extent from the object plane to the image plane, representing the sum of the object distance, the TTL, and any intermediate optical intervals.”

Furthermore, we agree that "tabletop-grade throughput" lack precision. We have therefore replaced it with *"matching the neuron yield of benchtop one-photon mesoscopes."*

We have revised the abstract as follows: " Airscope's optical design achieves a 2.84× shorter total track length than prior miniaturized microscopes while matching the neuron yield of benchtop

one-photon mesoscopes; a tailored wireless link supports wire-free calcium imaging from >8,500 neurons per animal (25-fold increase).”

Minor Comment 2 — "billions of neurons" without species context

Line 52, "billions of neurons": this seems to refer to larger brains, such as the human brain, but is mentioned without specifying this. That is confusing since the manuscript is about mouse brains, which contain fewer neurons.

Response: We thank the reviewer for catching this ambiguity. The sentence has been revised to specify: "*Understanding how brainwide neurons collectively produce these functions remains a grand scientific challenge, one requiring observation of neural circuits at the resolution of individual neurons.*" This provides appropriate context for both species.

Minor Comment 3 — Grammar and cost clarification (lines 69, 128, 269)

Line 69, "Even" should probably be "Even though". The cost of 2 dollars mentioned on line 128, is this only for the lens array or for the entire device? It must be for the lenses only, since the sensor and chips likely cost more - please clarify this. Line 269, "monthly" should probably say "month-long".

Response: We thank the reviewer for these careful corrections. All three have been addressed:

- Line 69: Revised to "*Even though...*"
- Line 128: Clarified to "*the elements are mass-produced by injection molding at <\$2 per module (Supplementary Table 3).*" A detailed per-element cost can be found in Supplementary Table 3.
- Line 269: Revised to "*month-long.*"

Additionally, to meet the journal's length restrictions, we have substantially rewritten the manuscript, and the sentences containing the flagged wording were removed or rephrased in the process. We have confirmed that the wording issues raised in the reviewer's inspection are no longer present in the revised text.

Minor Comment 4 — "to our knowledge" insertion and unlabeled axes

Line 445, "to our knowledge" appears to have been inserted in error. Multiple axes and color bars are unlabeled and only defined in the legend, for example Fig. 6m and 6r. There is space to add labels and it would make the figures easier to interpret. The same comment applies to a number of other figure panels.

Response: We thank the reviewer for these observations. We have:

- Removed the erroneous "to our knowledge" from line 445.
- Conducted a systematic review of all figure panels and added axis labels and color bar annotations directly to every panel where they were previously defined only in the legend. Specific panels that have been updated include original Fig. 6m, 6r (Now Extended Data Fig. 10c, i), and all other panels identified during our review. We believe the figures are now substantially easier to interpret independently.

Minor Comment 5 — Accessing code error

Referee #2 (Remarks on code availability): The GitHub address didn't work - I am guessing the authors forgot to make the repository public

Response: We thank the reviewer for their diligence in attempting to access our resources. Previously, the codebase was hosted via a cloud-storage repository for initial review and included as a supplementary software package.

For this revision, the complete repository—comprising hardware design documentation, data processing pipelines, and source code—has been transitioned to a fully public GitHub repository. The repository is now accessible at https://github.com/Neurallabware/Airscope_toolbox, ensuring full transparency and reproducibility in accordance with journal standards.

Reference

- 1 Meyer, A. F., Poort, J., O'Keefe, J., Sahani, M. & Linden, J. F. A Head-Mounted Camera System Integrates Detailed Behavioral Monitoring with Multichannel Electrophysiology in Freely Moving Mice. *Neuron* **100**, 46-60.e7 (2018).
- 2 Parker, P. R. L., Abe, E. T. T., Leonard, E. S. P., Martins, D. M. & Niell, C. M. Joint coding of visual input and eye/head position in V1 of freely moving mice. *Neuron* **110**, 3897–3906.e5 (2022).
- 3 Holmgren, C. D. *et al.* Visual pursuit behavior in mice maintains the pursued prey on the retinal region with least optic flow. *elife* **10**, e70838 (2021).
- 4 Padilla-Coreano, N. *et al.* Cortical ensembles orchestrate social competition through hypothalamic outputs. *Nature* **603**, 667–671 (2022).
- 5 Zong, W. *et al.* Large-scale two-photon calcium imaging in freely moving mice. *Cell* **185**, 1240–1256.e30 (2022).
- 6 Kingsbury, L. *et al.* Correlated Neural Activity and Encoding of Behavior across Brains of Socially Interacting Animals. *Cell* **178**, 429–446.e16 (2019).
- 7 Scherrer, J. R., Lynch, G. F., Zhang, J. J. & Fee, M. S. An optical design enabling lightweight and large field-of-view head-mounted microscopes. *Nat Methods* **20**, 546–549 (2023).
- 8 Chen, T. W. *et al.* Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295–300 (2013).
- 9 Zong, W. *et al.* Fast high-resolution miniature two-photon microscopy for brain imaging in freely behaving mice. *Nat Methods* **14**, 713–719 (2017).
- 10 Xue, F. *et al.* Multi-region calcium imaging in freely behaving mice with ultra-compact head-mounted fluorescence microscopes. *Natl Sci Rev* **11**, nwad294 (2024).
- 11 Jacob, A. D. *et al.* A Compact Head-Mounted Endoscope for In Vivo Calcium Imaging in Freely Behaving Mice. *Curr Protoc Neurosci* **84**, e51 (2018).
- 12 Ghosh, K. K. *et al.* Miniaturized integration of a fluorescence microscope. *Nat Methods* **8**, 871–878 (2011).
- 13 Cai, D. J. *et al.* A shared neural ensemble links distinct contextual memories encoded close in time. *Nature* **534**, 115–118 (2016).
- 14 Information on the (previous Version 3) Miniscope platform, <[http://miniscope.org/index.php/Information_on_the_\(previous_Version_3\)_Miniscope_platform](http://miniscope.org/index.php/Information_on_the_(previous_Version_3)_Miniscope_platform)> (2018).
- 15 Shuman, T. *et al.* Breakdown of spatial coding and interneuron synchronization in epileptic mice. *Nat Neurosci* **23**, 229–238 (2020).
- 16 Bollimunta, A. *et al.* Head-mounted microendoscopic calcium imaging in dorsal premotor cortex of behaving rhesus macaque. *Cell Rep* **35**, 109239 (2021).
- 17 de Groot, A. *et al.* NINscope, a versatile miniscope for multi-region circuit investigations. *elife* **9**, e49987 (2020).
- 18 Liberti, W. A., Perkins, L. N., Leman, D. P. & Gardner, T. J. An open source, wireless capable miniature microscope system. *J Neural Eng* **14**, 045001 (2017).
- 19 Barbera, G., Liang, B., Zhang, L., Li, Y. & Lin, D. T. A wireless miniScope for deep brain imaging in freely moving mice. *Journal of Neuroscience Methods* **323**, 56–60 (2019).
- 20 Miniscope V4, <http://miniscope.org/index.php/Miniscope_V4> (2020).

- 21 Guo, C. *et al.* Miniscope-LFOV: A large-field-of-view, single-cell-resolution, miniature microscope for wired and wire-free imaging of neural dynamics in freely behaving animals. *Sci Adv* **9**, eadg3918 (2023).
- 22 Zhao, P. P. *et al.* MiniXL: An open-source, large field-of-view epifluorescence miniscope enabling single-cell resolution and multi-region imaging in mice. *Sci Adv* **11**, eads4995 (2025).
- 23 Park, J., Brady, D. J., Zheng, G., Tian, L. & Gao, L. Review of bio-optical imaging systems with a high space-bandwidth product. *Advanced Photonics* **3**, 044001 (2021).
- 24 Zhang, Y. *et al.* Rapid detection of neurons in widefield calcium imaging datasets after training with synthetic data. *Nat Methods* **20**, 747-754 (2023).
- 25 Stringer, C. *et al.* Extracting large-scale neural activity with Suite2p. *bioRxiv*, 2026.2002.2004.703741 (2026).
- 26 Song, A., Gauthier, J. L., Pillow, J. W., Tank, D. W. & Charles, A. S. Neural anatomy and optical microscopy (NAOMi) simulation for evaluating calcium imaging methods. *Journal of Neuroscience Methods* **358**, 109173 (2021).
- 27 Carlo, C. N. & Stevens, C. F. Structural uniformity of neocortex, revisited. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 1488-1493 (2013).
- 28 Ebrahimi, S. *et al.* Emergent reliability in sensory cortical coding and inter-area communication. *Nature* **605**, 713-721 (2022).
- 29 Zhang, Y. L. *et al.* Long-term mesoscale imaging of 3D intercellular dynamics across a mammalian organ. *Cell* **187**, 6104–6122.e25 (2024).
- 30 Musall, S., Kaufman, M. T., Juavinett, A. L., Gluf, S. & Churchland, A. K. Single-trial neural dynamics are dominated by richly varied movements. *Nat Neurosci* **22**, 1677-1686 (2019).
- 31 Stringer, C. *et al.* Spontaneous behaviors drive multidimensional, brainwide activity. *Science* **364**, 255 (2019).
- 32 Lee, H. E. *et al.* Micro Light-Emitting Diodes for Display and Flexible Biomedical Applications. *Adv. Funct. Mater.* **29**, 1808075 (2019)
- 33 Kim, T. I. *et al.* Injectable, cellular-scale optoelectronics with applications for wireless optogenetics. *Science* **340**, 211-216 (2013).
- 34 Mesraoua, B. *et al.* Drug-resistant epilepsy: definition, pathophysiology, and management. *J. Neurol. Sci.* **452**, 120766 (2023).
- 35 Li, M. C. H. & Cook, M. J. Deep brain stimulation for drug-resistant epilepsy. *Epilepsia* **59**, 273–290 (2018).
- 36 Zheng, L. *et al.* Therapeutic potential of cAMP-mediated lysosomal pH modulation in ATP6V1B2-related neuropathology. *Cell Death Discov.* **12**, 199 (2026).
- 37 Fei, F. *et al.* Discrete subicular circuits control generalization of hippocampal seizures. *Nat. Commun.* **13**, 5010 (2022).
- 38 Kitanishi, T., Umaba, R. & Mizuseki, K. Robust information routing by dorsal subiculum neurons. *Sci. Adv.* **7**, eabf1913 (2021).

Nature-2025-10-27104 - Response to Reviewers' Comments

Response to all Reviewers

We sincerely thank the editor and all four reviewers for their positive evaluation of our revised manuscript and for their constructive comments. Reviewers 1 and 2 raised no further concerns. Reviewer 3 asked us to align the language on spinal cord imaging, causal interpretation, statistical inference and the terminology for the recorded cortical regions more closely with the experimental evidence, a point also raised in the editor's comment 2, and confirmed that no further experiments are needed once the spinal cord claims are restricted to population-level imaging. Reviewer 4 requested clarifications on data compression, multi-animal wireless recording, thermal load, illumination, behavioral tracking and behavior-tuned neurons, together with several minor additions. We have addressed all comments through revisions of the main text, Methods, figure legends and Supplementary Information, without new experiments. We summarize these revisions in the following table:

Revision	Requested by	Description	Related to
Level of resolution in spinal cord recordings	Reviewer 3; Editor	Spinal recordings are described as spatially resolved population activity (spinal components); the Results state explicitly that individual spinal neurons were not segmented, and single-neuron resolution is claimed only for the cortex.	Results; Discussion; legends of Fig. 6 and Supplementary Fig. 17; Supplementary Video 14
Scope of causal interpretation	Reviewer 3	The transection result is restricted to the stimulus-evoked ascending pathway, and Granger-causal influence is distinguished from anatomical or interventional causation.	Results; Methods
Details of the Granger causality analysis	Reviewer 3	Analysis window, treatment of cue and shock epochs, trial handling, covariates and stationarity assessment are specified.	Methods
Animal-level statistical inference	Reviewer 3	Statistical inference in the transection experiment is now based on animal-level analyses ($n = 3$ mice), including the fraction of stimulus-responsive neurons and the mean stimulus-evoked response. The neuron-level comparison is retained for visualization but is no longer used for statistical inference.	Supplementary Fig. 17p, q; Fig. R1

Revision	Requested by	Description	Related to
Terminology for the recorded regions	Reviewer 3	“Whole-brain”, “brain-wide” and “brainwide” are replaced with “dorsal cortex-wide”.	Results; Methods; legends of Fig. 6, Extended Data Figs. 9 and 10, Supplementary Figs. 14 and 17; Supplementary Video 12
Data compression and multi-animal wireless recording	Reviewer 4	New Methods paragraphs describe on-sensor JPEG compression, frame streaming and timestamping, and simultaneous recording from multiple Aircscopes.	Results; Methods; Supplementary Fig. 4
Thermal load and LED illumination	Reviewer 4	Power budget and heat dissipation are described, and illumination without collimation optics is clarified.	Methods
SAM2Mice and UDMT	Reviewer 4	Differences in output, training and initialization between SAM2Mice and UDMT are explained in our response.	Fig. R2
Behavior-tuned neurons in the tube test	Reviewer 4	Fractions of push-, resist- and retreat-tuned neurons and the number of neurons remaining after their removal are reported.	Fig. 4o; Extended Data Fig. 8l–n
References, comparison table and figure details	Reviewer 4	DeepInMiniscope is cited; Supplementary Table 1 indicates 1P, 2P or 3P excitation and includes two multiplexed multiphoton miniscopes; recording time (Fig. 1m–o), axis values (Supplementary Fig. 4d) and the SD extension board (Supplementary Fig. 5) are added.	Introduction; Supplementary Table 1; legends of Fig. 1 and Supplementary Figs. 4 and 5

Taken together, these text revisions address all remaining comments of Reviewers 3 and 4, including the editor’s request to tune down the claims on spinal cord recordings, without new experiments. The point-by-point responses follow.

Point-by-point response to reviewers' comments

Reviewer #1

Remarks to the Author

The authors have made substantial effort to comprehensively answer all of the reviewer comments. They agreed with the reviewers about optical characterization metrics, including the spatio-temporal bandwidth, provided all requested technical files and software for open access, adjusted the biological claims. The result is a substantially clearer paper with sharper quantifications and validation of the very impressive technical advance described here. I have no more concerns and thank the authors for the upfront discussion. Well done.

Response: We sincerely thank the reviewer for the careful evaluation over both rounds of review and for this encouraging assessment. We are grateful for the comments that motivated the revised optical characterization, including the spatio-temporal bandwidth product, and the open release of the technical files and software.

Reviewer #2

Remarks to the Author

The authors performed an extremely thorough and high-quality revision. My comments have been comprehensively addressed, and I have no further feedback apart from expressing my enthusiasm for this work.

Response: We are grateful to the reviewer for the constructive comments in the previous round and for the enthusiastic support of this work.

Remarks on code availability

I have inspected but not run the code, as this would be the appropriate prior infrastructure I do not currently possess.

Response: We thank the reviewer for inspecting the code. For readers without the acquisition hardware, the public data release provides raw recordings (figshare, <https://doi.org/10.6084/m9.figshare.33942097>), and the repository provides step-by-step Jupyter notebooks for processing and visualizing them (https://github.com/Neurallabware/Airscope_toolbox; archived at figshare, <https://doi.org/10.6084/m9.figshare.33942148>), with full documentation at <https://airscope-docs.readthedocs.io/en/latest/>.

Reviewer #3

Remarks to the Author

The manuscript presents an impressive technological advance. Airscope combines large field of view, low weight and wireless operation in a way that should enable experiments difficult with existing head-mounted approaches. The additional validation is extensive, and I do not see a need

for further experimental work except potentially on one point concerning the spinal cord recordings.

Response: We thank the reviewer for the positive assessment of Airscope and of the additional validation. Following the reviewer's recommendation, we have restricted the spinal cord claims to spatially resolved population imaging and revised the language accordingly, so that no additional experiments are required. Our point-by-point responses follow.

Major comment:

Major Comment 1 — Single-neuron resolution in spinal cord recordings

The manuscript convincingly establishes single neuron resolution in cortex through comparison with two-photon ground truth. However, I am less convinced that equivalent functional single-neuron resolution has been demonstrated in the spinal cord.

In the spinal experiments, the authors explicitly state that strong scattering necessitates the use of CellSort to extract four functional components rather than segmenting individual neurons. Figure 6 likewise describes these signals as four spinal cord components. The reported approximately 4 μm optical resolution is impressive, but optical resolution alone does not demonstrate that functional calcium signals can be reliably assigned to individual spinal neurons during freely moving recordings.

As currently presented, I would therefore interpret these data as spatially resolved spinal population or ensemble activity rather than validated single-neuron imaging. This distinction is important because spinal imaging is presented as a major capability of the platform. The authors need to be more explicit about single neurons resolution capabilities in the spinal cord or lack thereof.

If the authors intend to claim population level spinal imaging, I do not think further experiments are necessary, but the terminology should be revised accordingly. If, however, single neuron functional spinal imaging is intended as a central claim, I think direct experimental validation would be required, for example through sparse labelling or an independent cellular-resolution ground truth.

Response: We thank the reviewer for this important distinction, and we fully agree. Fourier ring correlation confirmed $\sim 4 \mu\text{m}$ optical resolution in spinal vascular images (Fig. 6c), but optical resolution alone does not demonstrate that functional calcium signals can be assigned to individual spinal neurons during free movement. Because strong scattering in spinal tissue required us to extract functional components with CellSort rather than to segment individual neurons, we do not claim single-neuron functional imaging in the spinal cord. As recommended by the reviewer, and in line with the editor's request to tune down this claim, we now present the spinal data as spatially resolved population imaging and define the level of resolution explicitly. Specifically:

1. Explicit definition of the level of resolution. The sentence describing the extraction of spinal signals in the Results now reads:

“Because spinal tissue scatters strongly, we extracted four functional components with CellSort rather than segmenting individual spinal neurons (Fig. 6f).”

2. Consistent terminology. “Spinal units”, “spinal ensembles” and “spinal cord neurons” have been replaced with “spinal components” or “spinal population activity” in the Results, in the

legends of Fig. 6f and k and Supplementary Fig. 17e, f and i, and in the description of Supplementary Video 14; in the legends of Fig. 6h and j and Supplementary Fig. 17g, which describe cortical and spinal responses jointly, “neuronal” has been replaced with “neural”. For example, the legend of Fig. 6f now reads “4 spinal components extracted via CellSort and activities”, and that of Fig. 6k reads “Response amplitudes of spinal cord components following auditory cue stimulation”.

3. Removal of cellular-resolution wording for the spinal cord. In the Results, “concurrent cellular-resolution monitoring of cortex and spinal cord in unrestrained animals has remained technically challenging” now reads “concurrent monitoring of cortex and spinal cord in unrestrained animals has remained technically challenging”, and “Airscope simultaneously recorded ~4,000 neurons across the right cortical hemisphere along with segments T13 and L1 of the spinal cord” now reads “Airscope simultaneously recorded ~4,000 cortical neurons across the right hemisphere together with population activity in spinal segments T13 and L1”. In the Discussion, “concurrent mesoscopic neuronal imaging of cortex and spinal cord” now reads “concurrent mesoscopic imaging of cortex and spinal cord”.

With these revisions, single-neuron resolution is claimed only in the cortex, where it is validated against two-photon ground truth (Extended Data Fig. 5), whereas the spinal recordings are presented as spatially resolved population activity.

Minor comments:

Minor Comment 1 — Causal interpretation of the transection experiment

1. The new transection experiment substantially strengthens the manuscript. Cortical responses to hindpaw stimulation are reduced almost to baseline after complete spinal transection.

This convincingly demonstrates that intact ascending spinal transmission is necessary for the cortical response to hindpaw stimulation. It does not, however, establish that the specific spinal components recorded by Airscope causally drive cortical activity, nor does it directly establish a spinal mechanism for the learned auditory cue response during fear conditioning. I would therefore keep the causal interpretation tightly restricted to the shock-evoked ascending pathway.

Response: We thank the reviewer for this positive assessment of the transection experiment and agree with this interpretation. We have restricted the causal interpretation to the stimulus-evoked ascending pathway: the clause stating that the transection provided “interventional support for the spinal-leading influence identified by the Granger analysis” has been deleted from the Results:

“Complete spinal-cord transection reduced trial-averaged neuron response amplitude and the fraction of stimulus-responsive cortical neurons to near zero (Supplementary Fig. 17l-q), indicating that intact ascending spinal transmission is necessary for cortical responses to hindpaw stimulation.”

Minor Comment 2 — Details of the Granger causality analysis

2. The revised term “Granger-causal influence” is appropriate, provided it remains clearly distinguished from anatomical or interventional causation. I would ask the authors to clarify the analysis window, whether cue and shock epochs were treated separately, how trials were

concatenated, whether movement or stimulus timing were included as covariates, and how stationarity was assessed. No additional experiment is required.

Response: We thank the reviewer for this suggestion. We have revised the Methods to state explicitly that Granger-causal influence denotes directed statistical dependence rather than anatomical or interventional causation, and we have added each of the requested analysis details. The revised Methods paragraph reads:

“Granger causality analysis. To characterize directed temporal dependencies among cortical regions and from the spinal cord to the cortex, we used the multivariate Granger causality toolbox (MVGC, version mvgc_v1.2; <https://github.com/SacklerCentre/MVGC1/>). A Granger-causal influence from X to Y indicates that the past activity of X improves the prediction of Y beyond the past activity of Y itself; it reflects directed statistical dependence and does not establish anatomical connectivity or interventional causation. Each trial was analyzed from 1 s before cue onset to 4 s after shock onset, and cue and shock epochs were analyzed within this single window. For each recording, the calcium traces were z-scored over time. Activity was averaged across the selected neurons within each cortical region, while spine activity was averaged across recorded spines. The resulting MVGC input had the dimensions variables × time points × trials. A total of 59 trials were pooled along the third, trial dimension. Neither movement measurements nor explicit cue/shock timing regressors were included as covariates. Stimulus timing was controlled by event alignment. The VAR model was estimated using the Morf LWR algorithm. The model order was selected by the Akaike information criterion (maximum 20). Pairwise-conditional Granger causality was computed from the fitted model and tested with the F-statistic, and connections with $p < 0.001$ after Benjamini–Hochberg false-discovery-rate correction were considered significant.”

In the Results, these findings are described as spinal-leading temporal coupling and Granger-causal influence, and the causal interpretation of the transection experiment has been restricted as described in our response to Minor Comment 1.

Minor Comment 3 — Animal as the experimental unit

3. The effect is large and convincing, but the reported significance for single-neuron responses uses >4k neurons from only three animals. Since neurons within an animal are not independent biological replicates, statistical inference should preferably use the animal as the experimental unit or an appropriate hierarchical analysis.

Response: We thank the reviewer for this important point and agree that neurons recorded within the same animal are not independent biological replicates. Indeed, the animal-level comparison for this experiment is shown in Supplementary Fig. 17q, in which the fraction of stimulus-responsive cortical neurons is compared within each of the three mice before and after transection (paired t-test, $p = 0.0369$, $n = 3$ mice). In this revision, we additionally quantified the mean stimulus-evoked response for each animal before and after transection (**Fig. R1**). This analysis, consistent with previous analysis, showed a significant reduction after transection (one-sided paired t-test, $p = 0.0358$, $n = 3$ mice), further supporting the effect at the animal level. Statistical inference is therefore based on these animal-level tests; the neuron-level comparison in panel p is retained for visualization only, and its significance marker has been removed. The revised legend of Supplementary Fig. 17p reads:

“p, Bar plot of single-neuron trial-averaged responses in control and spinal-transected mice (4,772 registered neurons from $n = 3$ mice; mean $\pm$ SEM).”

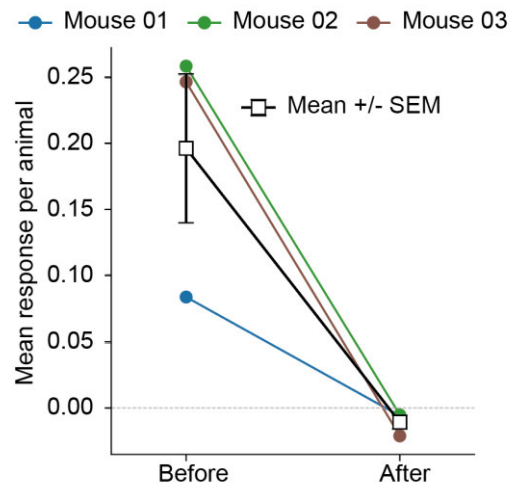


Fig. R1. Animal-level mean stimulus-evoked responses before and after transection. Lines represent individual animals and black squares indicate mean $\pm$ SEM (one-sided paired t-test, $*p=0.0358$, $n=3$ animals).

Minor Comment 4 — “Whole-brain” and “spinal neurons” terminology

4. The cortical recordings primarily sample layer 2/3 across dorsal cortex rather than the whole brain. I would therefore favour “dorsal cortex-wide” or “hemisphere-wide cortical” rather than “whole-brain.” Similarly, unless cellular identity is directly demonstrated in the spinal preparation, terms such as “spinal population activity”, “functional components” or “ensembles” would be more precise than “spinal neurons.”

Response: We thank the reviewer for this suggestion and agree that the cortical recordings primarily sample layer 2/3 of the dorsal cortex. We have replaced “whole-brain” and “brain-wide” with “dorsal cortex-wide” throughout the manuscript, and “spinal neurons” and related terms with “spinal population activity” or “spinal components” (see our response to the Major Comment). The revised instances are:

- Results: “Brain-wide recordings from thousands of neurons in interacting animals” now reads “*Dorsal cortex-wide recordings from thousands of neurons in interacting animals*”, and “Across the brain-wide population” now reads “Across the dorsal cortex-wide population”.
- Fig. 6g legend: “Brain-wide neuronal dynamics ... overlaid with the first principal component of whole-brain activity” now reads “*Dorsal cortex-wide neuronal dynamics ... overlaid with the first principal component of cortical activity*”.
- Legends of Extended Data Fig. 9k, Extended Data Fig. 10d, Supplementary Fig. 14g and Supplementary Fig. 17h, the description of Supplementary Video 12, and the Methods on orofacial behaviors: “whole-brain”, “brain-wide” and “brainwide” now read “*dorsal cortex-wide*”.

Recommendation

I would be comfortable with the manuscript moving forward once the language is adjusted as suggested to more faithfully reflect the experimental evidence. The only point that may require further experimental validation is the claim of single-neuron functional imaging in the spinal cord. If that claim is restricted to spatially resolved population imaging, and this is made clear in the manuscript, I do not think additional experiments are necessary.

Response: We thank the reviewer for this clear recommendation. As detailed above, we have restricted the spinal cord claims to spatially resolved population imaging and adjusted the language on causality, statistical inference and recorded regions to reflect the experimental evidence.

Reviewer #4

Remarks to the Author

The manuscript by Y. Zhang et. al. entitled “Deciphering cortex-wide neural dynamics of naturally behaving mice by 1-gram wireless mesoscope” reports a miniaturized microscope (termed as Airscope) that can be head-mounted on the mouse head to record its brain activity while the animal freely-behaves. The Airscope can be controlled wirelessly and transmit the imaging data wirelessly, which allows untethered imaging of mouse cortex and spinal cord. The Airscope weighs 1.0 gram, has 4-6 μm lateral resolution across a 6 mm field of view, and can perform recording at 12.5 Hz. Using the Airscope, the authors demonstrated neuronal activity recording under a variety of animal behaviors, including multi-animal free exploration over a 5 m² arena, competitive interaction between two animals, water reward searching in labyrinth, and response to pain by foot shocks and fear conditioning.

The results reported in the manuscript are very impressive. From the technical aspects, Airscope outperforms existing one-photon miniaturized microscopes, in particular those with wireless recording capability, in terms of spatial bandwidth product and spatiotemporal bandwidth product. In particular, the Chevalier architecture with few lenses design and the wireless transmission module with on-board data compression is quite innovative. The multi-mice tracking algorithm (SAM2Mice) also represents a very powerful tool. From the systems neuroscience perspective, Airscope enables a wide range of studies that were challenging to perform in the past. Large-scale calcium imaging can now be performed across multiple animals over a large arena without concerning the tethered wires tangling together. The biological findings, mostly about how neuronal activity encodes a specific animal behavior, are quite interesting. The compact footprint, light weight and wireless features of the Airscope should reduce the impact of the head-mounted device on the animals, and thus make the animal behavior more naturalistic. I believe Airscope represents a significant advancement in the field, and it can be widely adopted by the systems neuroscientists. The results presented in the paper are comprehensive and the statistical tests are appropriate. It has the merit to get published in Nature.

While the revised manuscript has addressed the previous reviewers’ comments and is of very high quality, it could be further strengthened by the consideration of the following comments.

Response: We thank the reviewer for the thorough evaluation of our manuscript, for recognizing the few-lens Chevalier-type optical design, the wireless module with on-board data compression and SAM2Mice as innovative, and for supporting publication in Nature. We address each comment below.

Main comments:

Comment 1 — Data compression and multi-animal wireless recording

1. One bottleneck of the previous wireless miniscopes is the limited transmission bandwidth. Airscope compresses the data prior to transmission. While there is a supplementary figure to describe this, the description is very sparse in the main text. In particular, the detail of the implementation is missing in Methods. More details on how to record the wireless data from multiple animals simultaneously should also be added in Methods.

Response: We thank the reviewer for this suggestion. We have added two paragraphs to the Methods that describe the implementation of on-board compression and the simultaneous recording of wireless data from multiple animals, and the corresponding sentence in the Results now refers to them (“*The compressed stream is delivered over Wi-Fi via bidirectional HTTP-GET (Fig. 1f, Supplementary Fig. 4 and Methods)*”). The new Methods paragraphs read:

“Data compression and wireless transmission. Each frame is compressed on the OV2640 image sensor by its built-in JPEG encoder, which applies a discrete cosine transform, quantization and entropy coding (Supplementary Fig. 4b). At $1,600 \times 1,200$ pixels and the encoder quality setting of the released firmware (6 on a scale of 0–63, on which lower values give higher image quality), compression reduces the data rate ~25-fold, from $\sim 30 \text{ MB s}^{-1}$ to $\sim 1 \text{ MB s}^{-1}$ (Supplementary Fig. 4d), without significant changes in calcium event rate or signal-to-noise ratio (Supplementary Fig. 4e–g). The ESP32 microcontroller runs an embedded HTTP server that streams the compressed frames as a motion-JPEG sequence and attaches a hardware timestamp to each frame, and a two-frame buffer absorbs transient network congestion. Frames and their timestamps can also be written to the optional SD extension board.”

“Simultaneous recording from multiple animals. All Airscopes connect as clients to a single Wi-Fi access point, each with its own IP address. The host acquisition software discovers the devices on the local network, identifies each device by its MAC address, and controls live preview, acquisition parameters and the start and end of recording for each device over HTTP. At the start of each recording, every Airscope requests a reference time from the host computer over UDP and logs it together with the round-trip delay, and the timestamp of each subsequent frame is stored relative to this reference, so that recordings from different Airscopes are aligned to a common host clock. With five Airscopes operating simultaneously on one access point, each device sustained $1,196.73 \pm 97.44 \text{ KB s}^{-1}$ at 12.41 ± 0.13 frames per second, with a frame-drop rate of $0.21 \pm 0.20\%$ and an end-to-end latency of $26.49 \pm 8.52 \text{ ms}$ (mean $\pm$ SD, $n = 5$ devices; Supplementary Fig. 4h–k).”

Comment 2 — Thermal management

2. There is a lack of discussion about the thermal management of the device. In particular, what is the total heat generated of the device on the animal?

Response: We thank the reviewer for raising this point. During continuous imaging, the fully assembled Airscope draws $\sim 0.6 \text{ W}$ in total, shared by the image sensor ($\sim 0.1 \text{ W}$), the microcontroller with its integrated Wi-Fi radio, and the LEDs with their driver; this value therefore sets an upper bound on the heat generated on the animal.

We have added a description of the thermal load and of the features that limit heat transfer to the animal to the Methods (PCB integration):

“Thermal load. During continuous imaging, Airscope draws ~0.6 W, which sets an upper bound on its heat dissipation. The main PCB is mounted on top of the housing rather than in contact with the skull, the 3D-printed housing has low thermal conductivity, and an on-board temperature sensor (MAX31875) allows the device temperature to be monitored during recording. In intermittent mode, the LED and sensors are switched off and the Wi-Fi radio enters standby between acquisition epochs, further reducing heat generation.”

Comment 3 — Collimation of the LED illumination

3. Regarding the LED illumination, is there a collimation lens after the LED? This needs to be clarified.

Response: We thank the reviewer for this question. No collimation lens is used after the LEDs. To minimize weight and footprint, the four LEDs illuminate the tissue directly through the excitation filters at an oblique angle, and illumination uniformity across the 6-mm FOV was instead achieved by optimizing the LED angle, lateral separation and height above the focal plane (Supplementary Fig. 2a–e). We have clarified this in the Methods (LED illumination optimization):

“The LEDs illuminate the FOV directly through the excitation filters, without collimation or focusing optics; illumination uniformity was therefore optimized through the LED geometry, as follows.”

Comment 4 — Comparison of SAM2Mice with UDMT

4. The authors developed SAM2Mice for multi-animal tracking. The authors also previously developed an “unsupervised deep transfer learning method for multi-animal tracking (UDMT)” for a similar task. It will be good to comment on how SAM2Mice is compared with UDMT.

Response: We thank the reviewer for this suggestion. UDMT¹ and SAM2Mice are complementary. UDMT is an unsupervised tracker that needs no training annotations: after the user selects each animal in the first frame, it fine-tunes its network on that video and outputs each animal's position trajectory and identity. SAM2Mice combines a YOLO-based auto-prompt network with the SAM2 foundation model to segment and track mice zero-shot, without per-video training or manual initialization, and outputs a pixel-level mask for each animal in every frame.

This pixel-level output is what our analyses required: SAM2Mice masks were used to determine which mouse occupied the food port or the running wheel during competition (Methods). Such occupancy cannot be resolved from the single point per animal that UDMT provides, because a centroid carries no information about body extent or contact with an object. As illustrated in Fig. R2, an animal's centroid can remain outside the predefined ROI even when part of its body overlaps the ROI. SAM2Mice was therefore not merely an alternative tracker but the tool that made these behavioral measurements possible.

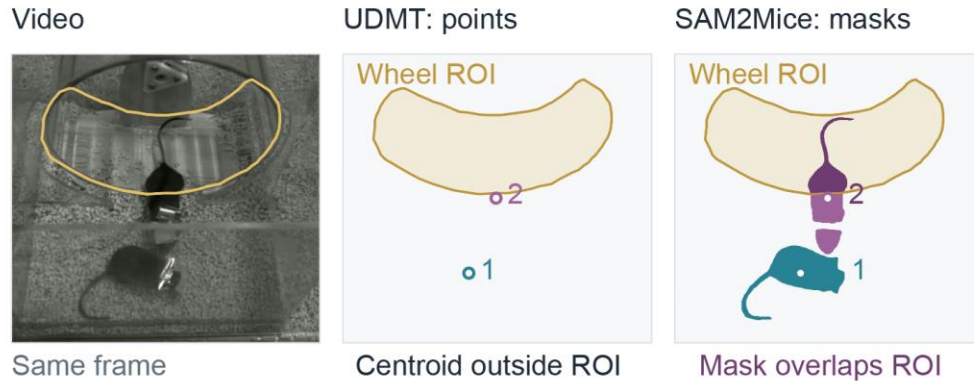


Fig. R2. Comparison of UDMT and SAM2Mice outputs. UDMT provides centroid-based trajectories, whereas SAM2Mice provides pixel-level masks that enable direct detection of animal–ROI overlap.

Comment 5 — Behavior-selective neurons and outcome decoding

5. For the competitive interaction experiments shown in Fig. 4, how many (or percentage of) neurons are selective for push, resist and retreat? When these neurons were removed from the decoder input, how many neurons were left? One interesting question is to evaluate how many neurons are needed to encode/predict the trial outcome. Such a question can only be answered through devices like Airscope. I would like to mention that such a result is interesting to see, but it is ok if the authors decide not to include in this manuscript, particularly when the results are already very dense in the manuscript.

Response: We thank the reviewer for this interesting question. The fractions of neurons responsive to push, resist and retreat are shown in Extended Data Fig. 8l. In dominant and subordinate animals, respectively, $4.04 \pm 0.17\%$ and $3.81 \pm 0.27\%$ of recorded neurons were responsive to push, $3.879 \pm 0.004\%$ and $3.50 \pm 0.41\%$ to resist, and $4.29 \pm 0.29\%$ and $3.93 \pm 0.26\%$ to retreat (mean $\pm$ SEM; high/low rank, $n = 14/7$, $2/5$ and $5/15$ mouse–pair sessions for push, resist and retreat, respectively; 8 mice in 2 cages). After all behavior-tuned neurons were removed, $3,114 \pm 210$ neurons per animal recording (94.1% retained on average) remained as decoder input, and decoding performance was not significantly reduced (Fig. 4o and Extended Data Fig. 8n).

We agree that determining how many neurons are needed to predict the trial outcome, for example with a neuron-dropping analysis, is an interesting question that recordings of this scale from interacting animals make possible. As the reviewer kindly noted, the results are already dense, and we have therefore not included this analysis in the present manuscript; we plan to address this question in future work.

Minor comments:

Comment 6 — Lensless approaches

6. Lensless approach is one competing technique to achieve large scale imaging with small footprint. The authors mentioned this category in the introduction. One most recent advance in this field is the following paper.

<https://www.science.org/doi/full/10.1126/sciadv.adr6687>

Response: We thank the reviewer for pointing us to this work. We now cite DeepInMiniscope² together with the other lensless approaches in the Introduction (“*Despite advances in multi-element lenses, low-NA designs, multiplexing, and lensless approaches, head-mounted devices trail tabletop mesoscopes ...*”).

Comment 7 — One-photon and multiphoton miniscopes in Supplementary Table 1

7. In Supplementary Table 1, the authors may want to add a column to indicate whether this is 1P or 2P/3P miniscope, or separate the table into two sessions for 1P, and 2P/3P. It will be more comprehensive to include the recently developed beam-multiplexed 2P/3P miniscope, which has an increased spatiotemporal bandwidth product, such as

[https://www.cell.com/cell-reports-methods/fulltext/S2667-2375\(25\)00257-7](https://www.cell.com/cell-reports-methods/fulltext/S2667-2375(25)00257-7)

<https://www.nature.com/articles/s41592-026-03125-7>

Response: We thank the reviewer for this suggestion. We have added a column to Supplementary Table 1 that indicates whether each system uses one-photon (1P), two-photon (2P) or three-photon (3P) excitation, and we have added the two recently reported multiplexed multiphoton miniscopes: the multiplexed miniaturized two-photon microscope (M-MINI2P)³, which increases the frame rate by imaging two FOVs simultaneously, and the head-mounted microscope combining two- and three-photon excitation⁴, which images five vertically separated planes across cortical layers in freely moving mice.

Comment 8 — Recording time in Fig. 1m–o

8. Fig. 1 m–o, please include the recording time in the caption.

Response: We thank the reviewer. Each open-field session lasted 15 min (Methods). We have added this information to the legend of panel m, which now reads “*m, Representative 15-min open-field trajectories ($40 \times 40 \text{ cm}^2$) without (Control) and with Airscope and battery; color, tortuosity; red square, $20 \times 20 \text{ cm}^2$ center.*”, and the legends of panels n and o now state that the metrics were computed from the same 15-min sessions.

Comment 9 — Vertical axis of Supplementary Fig. 4d

9. Supplementary Fig. 4d, quantitative value on the vertical axis is missing.

Response: We thank the reviewer for noticing this omission. We have added numerical tick labels to the vertical axis of Supplementary Fig. 4d (data throughput, MB s^{-1}) and state the values in the legend: “*d, Image data throughput before ($\sim 30 \text{ MB s}^{-1}$) and after ($\sim 1 \text{ MB s}^{-1}$) compression. Compression reduces the throughput more than 25-fold, below the Wi-Fi transmission limit of the microcontroller ($\sim 2.5 \text{ MB s}^{-1}$), which uncompressed data exceed.*”

Comment 10 — SD card option in Supplementary Fig. 5

10. Supplementary Fig. 5, does this assembly include the option of SD card placement?

Response: We thank the reviewer for this question. Yes, local storage is available as an option: the 0.31-g SD extension board plugs into the main PCB through a board-to-board connector

(Supplementary Fig. 3e, f and Methods). We have clarified this in the legend of Supplementary Fig. 5a: “*The optional SD extension board for local data logging (Supplementary Fig. 3e, f) plugs into the main PCB through a board-to-board connector before the Airscope is mounted.*”

Remarks on code availability

I went through the above github repo as well as the two other websites provided in the manuscript:

<https://airscope.org/devkit/>

<https://airscope-docxs.readthedocs.io/en/latest/>

The content is well organized and comprehensive. The Readme file is clear.

Response: We thank the reviewer for checking the repository and the online documentation, and for this positive feedback.

Reference

- 1 Li, Y. *et al.* Unsupervised transfer learning enables multi-animal tracking without training annotation. *Nat. Methods* **23**, 1058–1068 (2026).
- 2 Tian, F., Mattison, B. & Yang, W. DeepInMiniscope: deep learning–powered physics-informed integrated miniscope. *Sci. Adv.* **11**, eadr6687 (2025).
- 3 Zhang, Z. *et al.* High-speed neural imaging with multiplexed miniaturized two-photon microscopy. *Cell Rep. Methods* **5**, 101221 (2025).
- 4 Klioutchnikov, A. *et al.* Simultaneous two- and three-photon multiplane imaging across cortical layers in freely moving mice. *Nat. Methods* **23**, 1437–1446 (2026).