

# Volumetric two-photon imaging of neurons using stereoscopy (vTwINS)

Alexander Song<sup>1,6</sup>, Adam S Charles<sup>2,6</sup>, Sue Ann Koay<sup>2</sup>, Jeff L Gauthier<sup>2</sup>, Stephan Y Thiberge<sup>2,3</sup>, Jonathan W Pillow<sup>2,4</sup> & David W Tank<sup>2,3,5</sup>

Two-photon laser scanning microscopy of calcium dynamics using fluorescent indicators is a widely used imaging method for large-scale recording of neural activity *in vivo*. Here, we introduce volumetric two-photon imaging of neurons using stereoscopy (vTwINS), a volumetric calcium imaging method that uses an elongated, V-shaped point spread function to image a 3D brain volume. Single neurons project to spatially displaced 'image pairs' in the resulting 2D image, and the separation distance between projections is proportional to depth in the volume. To demix the fluorescence time series of individual neurons, we introduce a modified orthogonal matching pursuit algorithm that also infers source locations within the 3D volume. We illustrated vTwINS by imaging neural population activity in the mouse primary visual cortex and hippocampus. Our results demonstrated that vTwINS provides an effective method for volumetric two-photon calcium imaging that increases the number of neurons recorded while maintaining a high frame rate.

Two-photon excitation laser scanning microscopy<sup>1</sup> (TPM) enables high-spatial-resolution optical imaging in highly scattering tissues such as the mammalian brain. When TPM is combined with genetically encoded calcium indicators<sup>2,3</sup> or synthetic indicators that label neural populations<sup>4</sup>, intracellular calcium dynamics can be measured across a population of cells, thus providing a method for large-scale recording of neural activity at cellular resolution<sup>4,5</sup>. Expanding the number of simultaneously recorded neurons increases the power of population analysis methods in studies of neural coding and dynamics. To increase the number of neurons recorded with two-photon calcium imaging, volumetric imaging methods, such as multiplane imaging<sup>6</sup>, random access fluorescence microscopy<sup>7–9</sup>, and ultrasound lens scanning<sup>10</sup>, have been developed.

In traditional TPM<sup>1</sup>, a diffraction-limited, high numerical aperture (high NA) point spread function (PSF) is raster-scanned across a single plane. Volume imaging can be performed by sequentially moving the focal plane (or sample) up or down

between each raster scan, and this pattern is repeated for each volume measurement. This method can be implemented with movable objectives, remote focusing<sup>11</sup>, or a liquid lens<sup>6</sup>. However, if the frame rate for single-plane imaging is  $N$  frames/s, and the number of planes imaged per volume is  $m$ , then the aggregate-volume frame rate is decreased to  $N/m$ . Many calcium indicators<sup>12</sup> have on-response kinetics below 0.1 s. To capture these dynamics, volume frame rates must remain close to 10 Hz. With current resonant-scanner-based TPM ( $N \approx 30$  Hz), this requirement suggests that only a relatively low number of planes ( $m = 3, 4$ ) can be used for multiplane volumetric imaging.

Elongation of the PSF of the focused excitation beam along the optical axis, by using either a low-NA Gaussian beam focus or Bessel beam methods<sup>13</sup>, can be used with raster scanning to form a projection image of a volume<sup>14</sup>. This methodology is useful in applications such as functional imaging of dendritic spines in sample volumes with sparse neural expression of the indicator<sup>15</sup>. However, in samples with dense expression, such as those encountered in large-scale recording of a neural population *in vivo*, extending a single PSF axially causes neurons at different depths to be superimposed. Information about depth in the sample of individual neurons is lost, and demixing of fluorescence signals from individual neurons is compromised if their images substantially overlap.

Our method addresses these limitations by splitting an elongated PSF into two excitation beams. These beams are spatially separated and angled inward to create a stereoscopic V-shaped PSF configuration (**Fig. 1a**). Raster scanning with this PSF produces a 2D projection image that preserves information about neural activity at different depths. We refer to this method as vTwINS. The intuition underlying vTwINS is straightforward: the soma of any neuron in the 3D volume makes two contributions to the 2D projection image: one soma-shaped image for each arm of the V-shaped PSF. The spatial offset between these two images is equal to the distance between the two arms of the V at the neuron's depth in the volume, thus resulting in short distances between deep neurons and longer distances between shallower neurons (**Fig. 1a**).

<sup>1</sup>Department of Physics, Princeton University, Princeton, New Jersey, USA. <sup>2</sup>Princeton Neuroscience Institute, Princeton University, Princeton, New Jersey, USA.

<sup>3</sup>Bezos Center for Neural Circuit Dynamics, Princeton University, Princeton, New Jersey, USA. <sup>4</sup>Department of Psychology, Princeton University, Princeton, New Jersey, USA. <sup>5</sup>Department of Molecular Biology, Princeton University, Princeton, New Jersey, USA. <sup>6</sup>These authors contributed equally to this work. Correspondence should be addressed to D.W.T. (dwtank@princeton.edu).

Although vTwINS ensures that all neurons have distinct ‘paired’ spatial profiles in the projection image, identifying which pair of image regions reflects a single neuron’s activity is ill conditioned on the basis of single images. Recent methods have solved this problem by leveraging the temporal statistics of neural activity across frames (for example, principal component analysis and independent component analysis<sup>16</sup> and constrained non-negative matrix factorization (CNMF)<sup>17,18</sup>). We describe an inference algorithm based on orthogonal matching pursuit that exploits both the expected shapes of neural spatial profiles (pairs of rings or disks displaced along the axis of the V-shaped PSF) and the sparseness of neural activity. The image-pair separation  $\Delta$  in a neuron’s identified spatial profile also determines its axial position via the relationship  $d = 0.5(\Delta - \Delta_{\min})/\tan(\theta)$ , where  $\Delta_{\min}$  is the minimum interbeam distance of the PSF, and  $\theta$  is the beam angle from the axial direction (Fig. 1a,b). Thus, the demixing algorithm provides both the neuron’s fluorescence time course and its location in the volume.

Here, we describe the optics developed to produce the vTwINS PSF and demonstrate images and image time series generated through this method. We then present the algorithm that we developed for identifying active neurons in these time series and demixing fluorescence transients. Finally, using the combined imaging system and algorithm, we demonstrate large-scale recording of

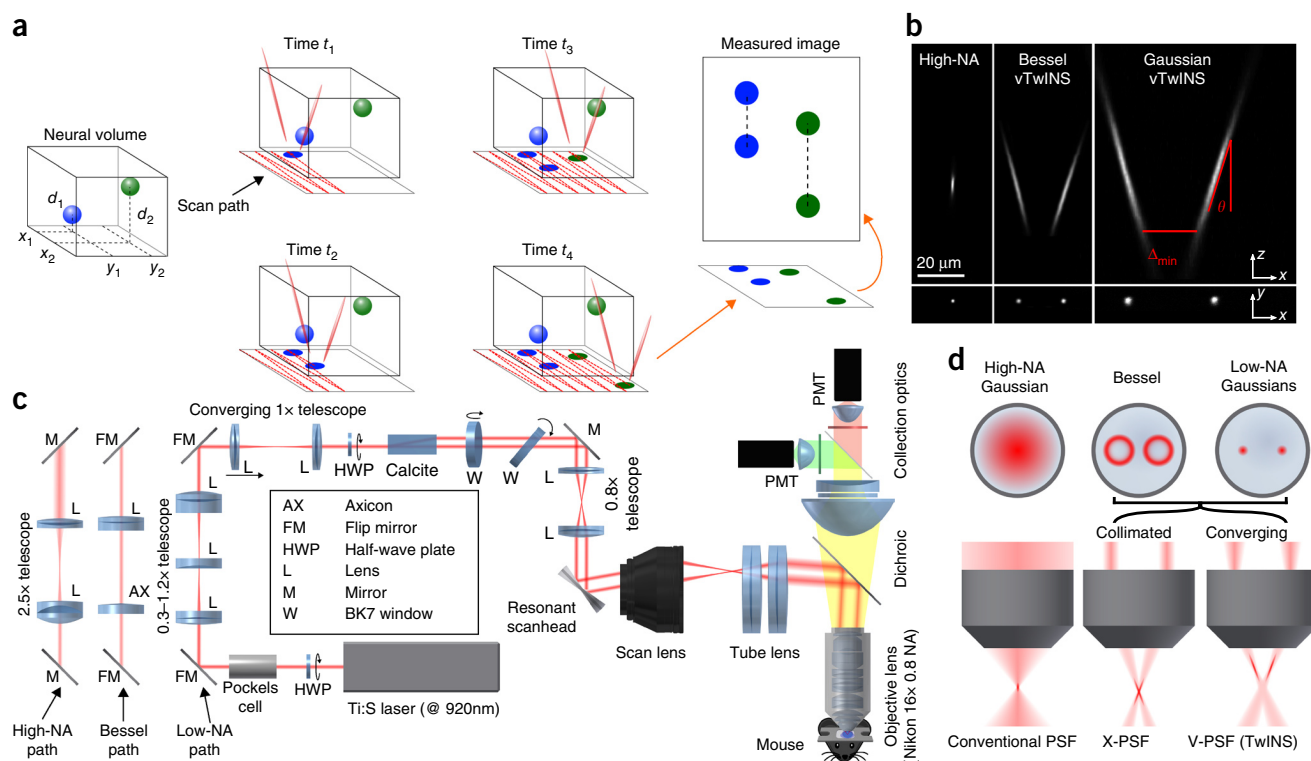
neurons expressing the genetically encoded calcium indicator GCaMP in the visual cortex and hippocampus in awake mice.

## RESULTS

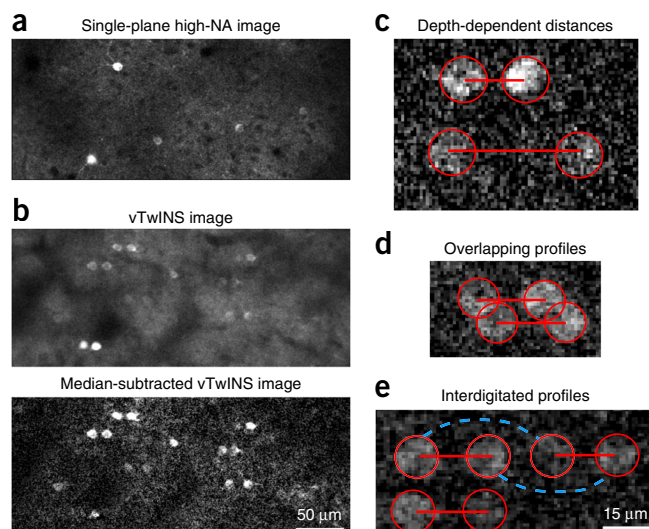
### vTwINS optics

In a vTwINS microscope, the diffraction-limited PSF of traditional TPM is replaced with an elongated V-shaped PSF produced from two intersecting Gaussian beams, or Bessel beams (Fig. 1b). We created the V-shaped PSF by dual-beam excitation through a single objective lens (Fig. 1c,d). To produce a V-shaped PSF, the back aperture is illuminated with a pair of small Gaussian beams or rings (Bessel beams<sup>19</sup>). Adjusting the separation distance, beam convergence, and beam parameters changes the angle, offset, and extension of the PSF, respectively (Online Methods, **Supplementary Fig. 1a** and **Supplementary Note 1**).

As an initial proof of principle that vTwINS can spatially localize objects in a 3D volume, we imaged fluorescent beads embedded in agarose (**Supplementary Fig. 2** and **Supplementary Note 2**). We localized the center positions of the beads to  $2.7 \pm 1.6 \mu\text{m}$  ( $N = 31$ ) against a reference image stack, conditions more accurate than necessary for imaging neuronal-cell bodies. When beads were localized over a large axial range, we found that *in vivo* (Online Methods) the effective axial illumination length of each elongated beam was approximately  $1/e$  full width of the maximum intensity.



**Figure 1** | vTwINS concept and design. **(a)** vTwINS uses a V-shaped PSF to image neural volumes. During scanning, the two PSF arms intersect neurons at different depths (for example, the blue and green stylized neurons) at different time intervals. Deep neurons intersect the second arm soon after the first. Shallow neurons take longer for the second arm to intersect. Each neuron thus appears twice, and the distance between images indicates depth. **(b)** Example PSFs for diffraction-limited (high-NA) TPM and vTwINS microscopes using Bessel and low-NA Gaussian beams. **(c)** The vTwINS microscope consists of a beam-shaping module and a conventional two-photon microscope. The three optical paths generate the PSFs shown in **b**. In the Bessel and Gaussian (low-NA) vTwINS paths, lenses adjust the PSF’s axial extent, and a birefringent block (calcite) splits the beam in two and sets the PSF angle. **(d)** The back-aperture illumination profiles for the three paths in **c**. In the high-NA (conventional TPM) path, the overfilled back aperture is focused to a point. In the Bessel and low-NA Gaussian paths, two beams are focused to form each arm of the PSF. The beam divergence is adjusted with the 1x telescope before the calcite block to separate the two arms of the X-PSF and to form the V-PSF.

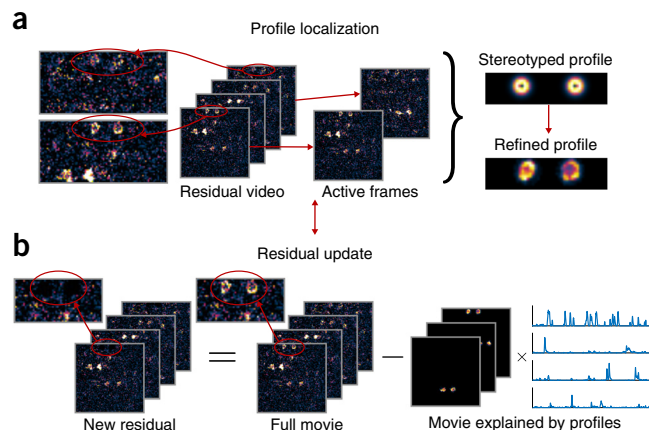


**Figure 2** | Example vTwINS images. All images are averages of five consecutive frames taken at 30 Hz. (a) Diffraction-limited TPM single-plane image of GCaMP6f in the mouse visual cortex. (b) vTwINS scan of the same V1 area as in a, showing paired somas of active neurons and a decreased SNR because the background levels are much higher. Subtracting the temporal median at each pixel highlights neural activity. (c) Two fluorescing neurons imaged by vTwINS at different depths have different distances between the image pairs. Red circles indicate the different images, and red lines connect corresponding image pairs. (d) vTwINS images typically have overlapping spatial profiles. (e) Neurons aligned in the direction parallel to the plane of the V-shaped PSF (which is the same as the fast scan direction in our implementation) can create ambiguity in the spatial-profile image-pair assignment. Both the solid red lines (the true pairing) and the dashed blue lines indicate realizable distance pairings corresponding to different neuron positions, and temporal activity must be used to resolve this ambiguity.

### vTwINS calcium imaging

In diffraction-limited TPM (Fig. 2a), a single soma-shaped spatial profile of high fluorescence intensity is observed when calcium transients are produced in an active neuron, and the cell soma of GCaMP-expressing quiescent cells can typically be resolved as well<sup>20</sup>. vTwINS images are qualitatively different. Active neurons are represented as two bright soma-shaped regions (disk or ring; Fig. 2b), and the images of quiescent neurons are typically not resolved, because the projection produces an increased and more uniform background intensity. The geometry of the vTwINS PSF decreases the signal of active neuropil and spreads it out over large regions, thus resulting in a broad, time-varying addition to the background intensity. When multiple cells are simultaneously active, many soma pairs become visible. Pairs from different cells have different spatial separations (Fig. 2c) representing different depths of the cell somas in the volume.

These properties of vTwINS-based calcium imaging data (Fig. 2) introduce unique challenges in demixing spatial profiles of neural activity to extract the fluorescence time traces of individual cells. First, there is a lower signal-to-noise ratio (SNR) per cell than that in diffraction-limited TPM, owing to the axially extended PSFs (Fig. 2b). Second, the spatial profiles of cells under vTwINS may partially overlap (Fig. 2d) and typically consist of disjoint regions. Whereas the geometry of vTwINS decreases the maximal overlap between profiles (Online Methods, Supplementary Note 3 and Supplementary Fig. 3), the disjoint nature of the



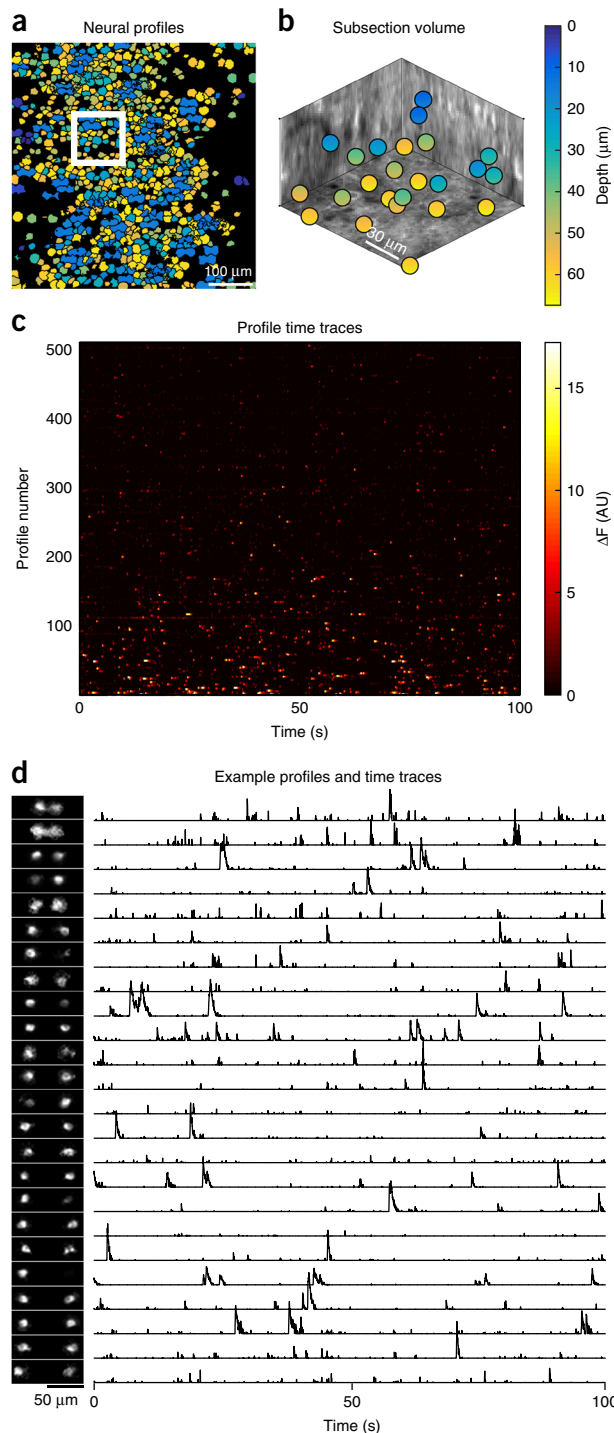
**Figure 3** | Sparse convolutional iterative shape matching (SCISM) for demixing vTwINS data. (a) Example stereotyped neuron image pairs with different distances are matched across frames to determine the most likely pair. The new profile is refined by locally masking and averaging frames closely aligned with the stereotyped spatial profile. (b) The new profile is added to the set of spatial profiles, and the time traces for all spatial profiles are calculated via non-negative LASSO. The residual movie is recomputed by subtraction of the contribution of the current set of spatial profiles (the sum of outer products of the spatial profiles and their time traces). The algorithm then finds the next spatial profile by restarting and operating on the new residual.

profiles violates the spatial locality assumption in current demixing methods<sup>16,18</sup> (Online Methods and Supplementary Fig. 4b). Third, neurons coaligned in the fast-scan direction may create ambiguous, interdigitated spatial-profile pairs (Fig. 2e). Finally, intensity differences between the two images in a pair may result from the nonuniform scattering between the two beam paths (for example, owing to varying tissue properties).

### vTwINS profile identification and demixing

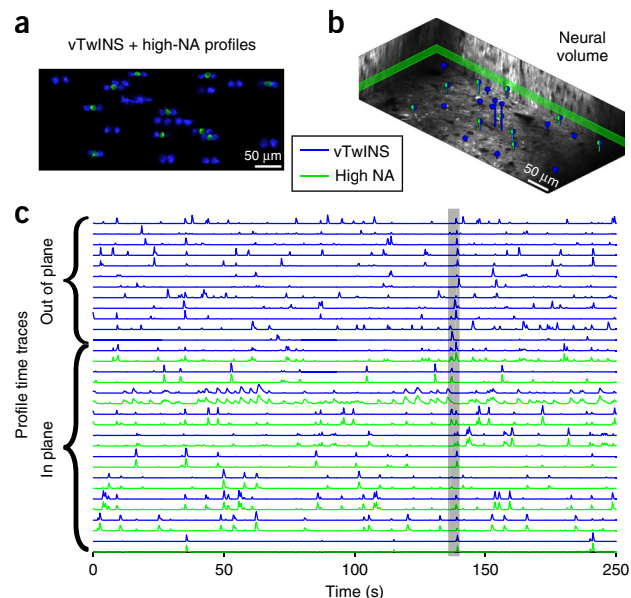
We addressed the challenges of analyzing vTwINS data with sparse convolutional iterative shape matching (SCISM), a demixing method that explicitly seeks horizontally separated image pairs (Online Methods, Fig. 3 and Supplementary Fig. 5). As a preprocessing step, we motion-corrected, temporally averaged, and spatially binned the raw-image time series (Online Methods). At each iteration, candidate spatial profiles, consisting of stereotyped profiles (pairs of annuli separated in the fast-scan direction with different separation distances), are compared with the measured fluorescence across the field of view (FOV) (Fig. 3a and Supplementary Fig. 5a,b). Simultaneously estimating pairs of projections increases the available signal to distinguish the neural profile from the noise fluctuations (Supplementary Fig. 6 and Supplementary Note 4). The stereotyped profile most correlated with the data is then selected (Supplementary Fig. 5c), and the most highly correlated images are used to refine the profile shape to better match the data (Supplementary Fig. 5d). This step allows SCISM to handle spatial-profile pairs in which one beam path has lower intensity. The new profile is added to the set of active spatial profiles, and the corresponding time traces are estimated by using non-negative LASSO<sup>21</sup> (Supplementary Fig. 5e). Finally, the data residual is calculated by subtracting out the component of the data captured by the current set of active spatial profiles (Fig. 3b and Supplementary Fig. 5f); this residual is used in the next recursion to determine the next spatial profile.





**Figure 4** | Demixed spatial profiles and calcium activity in the mouse visual cortex. (a) Full set of spatial profiles, color-coded by depth, showing substantial overlap. (b) 3D locations of the spatial profiles located in the white box in a. (c) Time traces of spatial profiles in a 100-s time interval. (d) Example subset of spatial profiles (chosen from the white box in a and sorted by depth) and corresponding normalized time traces. The increasing separation distance as a function of depth reflects the inverted V shape of the PSF used in this recording. F, fluorescence; AU, arbitrary units.

This procedure iteratively selects spatial profiles in a greedy manner in order of correlation strength with the data, by using both spatial and temporal statistics to determine the most likely spatial profile at each iteration. Specifically, SCISM leverages



**Figure 5** | Simultaneous imaging of the visual cortex with conventional two-photon microscopy (green) and vTwINS (blue). (a) Maximum projection of 5 s of activity for vTwINS (blue) and single-plane imaging (green). (b) Volumetric depiction of vTwINS extracted spatial-profile locations and depth. Green and blue profiles indicate locations of cells that were matched in single-plane activity. The single-plane slice is outlined in green. (c) Time traces corresponding to cells in a. The gray bar indicates the 5-s period of activity used to isolate cells.

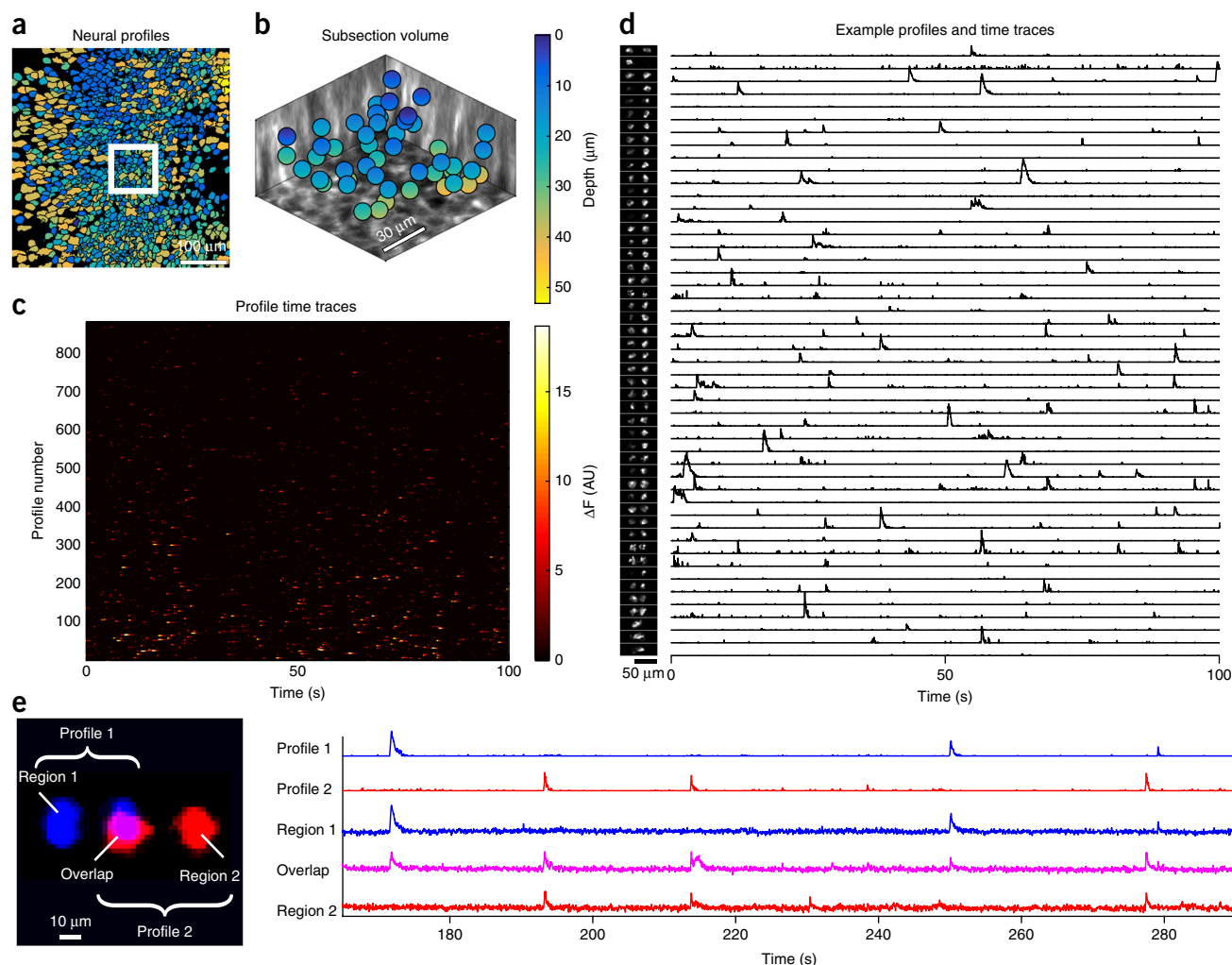
sparsity in neural activity as well as the spatial constraint that each spatial profile consists of two areas separated in the fast-scanning direction. Sparse neural activity is particularly important because it allows for minimal cross-contamination resulting from spatially overlapping neurons. After spatial profiles are determined with SCISM, full-resolution time-trace estimates are obtained by using nontemporally averaged data via non-negative LASSO.

### Large-scale recording in the mouse visual cortex

Head-restrained GCaMP6f-expressing transgenic mice, running on a spherical treadmill, were presented with a visual stimulus sequence consisting of randomly placed Gabor patches (Online Methods). We performed vTwINS imaging in layer 2/3 of the primary visual cortex (V1). We acquired images in a  $550\ \mu\text{m} \times 550\ \mu\text{m}$  area with a  $45\text{-}\mu\text{m}$ -long inverted-V PSF (full width at half maximum (FWHM),  $60\ \mu\text{m}$   $1/e$  full width) at a 30-Hz frame rate over a 14-min imaging session.

We preprocessed the time-series fluorescence data with rigid-motion correction and spatiotemporal averaging (Online Methods, **Supplementary Fig. 7a** and **Supplementary Videos 1** and **2**). The 511 spatial profiles obtained via SCISM (**Fig. 4**) showed substantial overlap, as expected from the high density of GCaMP-expressing cells and the vTwINS PSF. We used the vTwINS PSF to extract the 3D cell positions (Online Methods and **Fig. 4a,b**). The demixed spatial-profile activity traces (**Fig. 4c** and **Supplementary Fig. 8**) had the expected temporal statistics of sparsely firing neurons. Because SCISM is an iterative method that extracts highly active spatial profiles first, the time traces were ordered on the basis of how correlated the profiles were with the data.

The spatial-profile volumetric locations (**Fig. 4b**) indicated that vTwINS records activity across the entire axial extent of the PSF.



**Figure 6** | Demixed spatial profiles and calcium activity in the mouse hippocampus. (a) Full set of spatial profiles, color-coded by depth. (b) 3D locations of the spatial profiles from the white box in a. (c) Time traces of spatial profiles in a in a 100-s time interval. (d) Example subset of spatial profiles (chosen from the white box in a and sorted by depth) and corresponding normalized time traces. Notably, some profiles have sparse activity and do not contain transients in the displayed 100-s range. (e) Example demixed spatially overlapping profiles. Profile 1 (blue) and profile 2 (red) spatially overlap yet have demixed time traces (right). Averaged raw fluorescence traces from pixels in the overlapping region (overlap) are a linear combination of the traces from profile 1 (region 1) and profile 2 (region 2). The region 2 time trace also contains a transient from yet another profile at 230 s.

The range of axial depths captured by vTwINS was further illustrated by plotting the spatial profiles in a  $107\ \mu\text{m} \times 107\ \mu\text{m}$  subsection of the FOV (Fig. 4d), sorted by inferred depth (Fig. 4d) and the corresponding position in a 3D anatomical volume (Online Methods and Supplementary Fig. 9). The corresponding spatial-profile activity traces (Fig. 4d) also showed that cell transients were well isolated, despite the highly overlapping spatial profiles.

To validate that the neural activity recorded with vTwINS is comparable to that obtained through standard methods, we simultaneously imaged an entire neural volume with vTwINS and analyzed a single slice of the volume with diffraction-limited TPM. Both data sets were collected at 30 Hz over a  $470\ \mu\text{m} \times 200\ \mu\text{m}$  overlapping area (Supplementary Figs. 1b and 6c, and Supplementary Videos 3 and 4), whereas the vTwINS data set extended over the  $38\text{-}\mu\text{m}$ -long PSF (FWHM,  $52\ \mu\text{m}$  1/e full width). We demixed the vTwINS data with SCISM, and we extracted spatial profiles and activity traces from the single-plane data by using a modified CNMF algorithm<sup>22</sup> as an independent comparison (Online Methods).

Comparison of spatial profiles from the simultaneous recordings (Fig. 5) indicated that vTwINS captures both neural activity overlapping with the single-slice TPM data and activity at other depths. Overall, in a 10-min imaging session, we found 454 spatial profiles in the volume by using vTwINS, as compared with 169 spatial profiles found in the single-plane diffraction-limited data. When we restricted our analysis to a 5-s window for visualization purposes (Fig. 5a,b), activity traces corresponding to the spatial profiles of cells identified in both the single plane and the volume showed high correlation between the two imaging modalities. In the single-slice data set, 116 spatial profiles had  $>1$  transient per minute. Of these, 98 (84%) had a matching spatial profile in the vTwINS data (Supplementary Fig. 10). Of the remaining single-slice spatial profiles, many had very low SNR, thus suggesting that their activity fell below the vTwINS' lower SNR level. These correlations indicated that, as compared with TPM, vTwINS captures most of the activity at a particular given depth while also capturing additional activity elsewhere in the volume.

## Large-scale recording in the mouse hippocampus

In a more challenging application of vTwINS, we recorded and demixed activity from the CA1 region of the mouse hippocampus. In this region, neuronal-cell soma are densely packed in a well-defined layer, thus producing high spatial overlap in vTwINS data. To induce activity in CA1, we trained water-restricted mice to run down a linear track in a virtual reality system<sup>23</sup> to collect water rewards (Online Methods). We collected images over a 14-min session in a  $470\ \mu\text{m} \times 470\ \mu\text{m}$  area with a  $35\text{-}\mu\text{m}$  long vTwINS PSF (FWHM,  $45\ \mu\text{m}$   $1/e$  full width, noninverted V) at 30 Hz (Supplementary Fig. 7b and Supplementary Videos 5 and 6). We processed and analyzed CA1 recordings by using the same pre-processing and SCISM demixing as described for the V1 data.

We calculated the 3D positions for each of the 882 spatial profiles found through SCISM (Fig. 6a,b) and demixed their activity (Fig. 6c and Supplementary Fig. 11). The tendency for neurons toward the center of the FOV or neurons toward the edges of the FOV to be located in shallower or deeper locations, respectively, indicated that the vTwINS spatial profiles captured the curvature of CA1 (Fig. 6a). A subset of spatial profiles in a  $92\ \mu\text{m} \times 92\ \mu\text{m}$  section (Fig. 6b,d) of the FOV demonstrated that the inferred 3D locations matched well with the anatomical  $z$  stack (Supplementary Fig. 12).

Despite the highly overlapping spatial profiles resulting from the vTwINS PSF and high neural density, SCISM successfully demixed spatial profiles in CA1. Fluorescence time courses in different regions of two overlapping spatial profiles illustrated the demixed time traces (Fig. 6e). The trace from the overlapped region of the two cells contained transients from both nonoverlapping regions, whereas the demixed traces contained only the single-profile-region activity. Notably, one transient at 230 s in the trace of region 2 was missing from the overlap trace, thus indicating that this transient originated from a third profile and was successfully demixed in profile 2's time trace.

One caveat of vTwINS is the increased background from simultaneous scanning of the V-shaped PSF. As an alternative vTwINS implementation, we also explored in mouse CA1 a variation of vTwINS in which we alternated between each half of the vTwINS PSF on consecutive frames (Online Methods and Supplementary Fig. 1c). Illumination with a single beam ( $40\text{-}\mu\text{m}$ -long PSF; FWHM,  $56\ \mu\text{m}$   $1/e$  full width) increased the SNR per frame; however each beam can be recorded at only half the frame rate (15 Hz). By merging the 30-Hz interpolated recording, we used SCISM to locate spatial profiles separated by the width of the image (Supplementary Figs. 13a–c and 14). The temporal separation of the two channels decreased the total background neuropil, thereby improving the SNR of the recorded transients (Supplementary Fig. 13d–f).

## DISCUSSION

Early strategies for large-scale recording with calcium imaging generally used the spatial resolution of the optical instrumentation to ensure that the fluorescence from individual neurons formed spatially independent, disjoint sets. Spatial separation was the basis for manual selection of neural regions of interest, a method that has widely been used as a mask for extracting the time traces of individual cells. In practice, however, perfect separation of adjacent cell signals has been difficult to achieve for densely labeled cells. As a result, demixing algorithms<sup>16,17,22</sup> have been developed to identify spatial profiles by allowing individual pixels to have

contributions from multiple neurons. vTwINS (and also a recently described multiplane technique<sup>24</sup>) use this mixing assumption as a starting point for the development of the optical instrumentation. The vTwINS V-shaped PSF increases signal mixing in individual pixels but also ensures that each neuron has a unique spatial profile that can be efficiently used in the co-designed demixing algorithm to extract the time traces and 3D locations of individual cells. We anticipate that this strategy in which optical instrumentation and the demixing algorithm are co-designed for large-scale recording may be generalizable to other excitation geometries (for example, three beams and multiple objectives).

vTwINS confers the ability to seek specific spatial-profile shapes while maintaining the flexibility to adapt to variations in neuronal shape. SCISM permits the specification of these shapes as guides to locate relevant activity while still balancing the generally expected temporal statistics of neural activity. Current automated methods do not use such detailed spatial information, focusing instead on temporal demixing<sup>25–29</sup> with no spatial constraints<sup>16,17</sup> or on using generic locality assumptions (such as that spatial profiles must be fully contained in a constrained region)<sup>22,30</sup>. The ability of SCISM to adapt profiles to the data also differentiates it from standard matching-pursuit-style algorithms<sup>31–33</sup>, which assume a fixed dictionary of features. Although we designed SCISM to seek features specific to vTwINS imaging, it could easily accommodate other spatial-profile shapes in future imaging methods.

vTwINS, as compared with the most similar current approaches, offers a number of advantages. Unlike fast point-scanning strategies<sup>10</sup>, vTwINS has a substantially lower peak nonlinear excitation (Supplementary Note 5), thereby allowing for scanning in larger volumes. Additionally, the vTwINS excitation configuration is comparatively power efficient<sup>34</sup> and has high total two-photon excitation for a given average power. As compared with other multiplexed two-photon approaches<sup>24</sup>, vTwINS requires no additional recordings to initialize components and ensures the uniqueness of spatial profiles. Finally, as compared with volumetric TPM with single, elongated PSFs, the use of two beams in vTwINS decreases the ambiguity among spatial profiles (Supplementary Figs. 3 and 4a).

vTwINS, like other TPM techniques, is subject to practical concerns of imaging *in vivo*. For our recordings, motion correction was not compromised, owing to the elongated PSF, and sufficient high-spatial-frequency features in the imaged regions remained for accurate motion correction. Alternatively, in brain regions with low spatial structure, nuclear-localized RFP may be used for accurate motion correction. Nonlinear phototoxicity<sup>35,36</sup> in vTwINS is decreased (Supplementary Note 5) as compared with that in high-NA TPM, and heating<sup>37</sup> is limited to 100 mW per excitation beam.

Additional work should be able to further optimize vTwINS for other applications. vTwINS has additional background fluorescence from neuropil contamination, but it might work particularly well with a nuclear-localized GCaMP<sup>38</sup>, which would substantially decrease background neuropil fluorescence and improve the SNR. Furthermore, more flexible spatial information can be implemented in SCISM to improve demixing of neural signals from neuropil contamination. For brain regions with limited optical access (for example, the hippocampus<sup>23</sup> or medial entorhinal cortex<sup>39</sup>), smaller angles between arms or separation distance may be necessary. The choice between Bessel beams and Gaussian beams requires additional study. Bessel beams offer flexibility in controlling the axial profile and lateral resolution<sup>40</sup>. Gaussian



beams, although less flexible, are simpler to implement and have higher two-photon excitation (Online Methods, **Supplementary Fig. 1**). Finally, as a complementary method, vTwINS can be paired with existing improvements, such as sequential plane imaging (for example, remote focusing<sup>11</sup> or liquid lens<sup>6</sup>) to image larger volumes or to take advantage of the improved SNR arising from use of a lower-repetition-rate laser<sup>34</sup>.

## METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the [online version of the paper](#).

*Note: Any Supplementary Information and Source Data files are available in the online version of the paper.*

## ACKNOWLEDGMENTS

We thank C. Domnisoru, R. Low, and B. Scott for their insightful thoughts and comments. We also thank J. Homann for assistance in using the Psychophysics Toolbox. D.W.T. was supported by NIH grants R01MH083868 and U01NS09054, and the Simons Collaboration on the Global Brain (SCGB 328057). A.C. was supported by an NIH NRSA Training Grant in quantitative neuroscience (T32MH065214). J.W.P. was supported by grants from the McKnight Foundation, the Simons Collaboration on the Global Brain (SCGB 325407), and an NSF CAREER Award (IIS-1150186).

## AUTHOR CONTRIBUTIONS

D.W.T. conceived the project. A.S. and S.Y.T. designed and constructed the vTwINS microscope. S.A.K. and J.L.G. performed the surgery on the mice. A.S. trained the mice and performed the imaging experiments. A.S.C. and J.W.P. designed the SCISM algorithm. A.S.C. implemented SCISM and applied the method to the vTwINS data. A.S. and A.S.C. analyzed the results. A.S., A.S.C., J.W.P., and D.W.T. wrote the manuscript, and all authors provided comments and contributions. J.W.P. and D.W.T. supervised the project.

## COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

- Denk, W., Strickler, J.H. & Webb, W.W. Two-photon laser scanning fluorescence microscopy. *Science* **248**, 73–76 (1990).
- Mank, M. & Griesbeck, O. Genetically encoded calcium indicators. *Chem. Rev.* **108**, 1550–1564 (2008).
- Tian, L., Akerboom, J., Schreiner, E.R. & Looger, L.L. Neural activity imaging with genetically encoded calcium indicators. *Prog. Brain Res.* **196**, 79–94 (2012).
- Stosiek, C., Garaschuk, O., Holthoff, K. & Konnerth, A. *In vivo* two-photon calcium imaging of neuronal networks. *Proc. Natl. Acad. Sci. USA* **100**, 7319–7324 (2003).
- Dombeck, D.A., Khabbazi, A.N., Collman, F., Adelman, T.L. & Tank, D.W. Imaging large-scale neural activity with cellular resolution in awake, mobile mice. *Neuron* **56**, 43–57 (2007).
- Grewe, B.F., Voigt, F.F., van 't Hoff, M. & Helmchen, F. Fast two-layer two-photon imaging of neuronal cell populations using an electrically tunable lens. *Biomed. Opt. Express* **2**, 2035–2046 (2011).
- Göbel, W., Kampa, B.M. & Helmchen, F. Imaging cellular network dynamics in three dimensions using fast 3D laser scanning. *Nat. Methods* **4**, 73–79 (2007).
- Duemani Reddy, G., Kelleher, K., Fink, R. & Saggau, P. Three-dimensional random access multiphoton microscopy for functional imaging of neuronal activity. *Nat. Neurosci.* **11**, 713–720 (2008).
- Kirkby, P.A., Srinivas Nadella, K.M. & Silver, R.A. A compact Acousto-Optic Lens for 2D and 3D femtosecond based 2-photon microscopy. *Opt. Express* **18**, 13721–13745 (2010).
- Kong, L. *et al.* Continuous volumetric imaging via an optical phase-locked ultrasound lens. *Nat. Methods* **12**, 759–762 (2015).
- Botcherby, E.J., Juskaitis, R., Booth, M.J. & Wilson, T. An optical technique for remote focusing in microscopy. *Opt. Commun.* **281**, 880–887 (2008).
- Chen, T.-W. *et al.* Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295–300 (2013).
- Botcherby, E.J., Juskaitis, R. & Wilson, T. Scanning two photon fluorescence microscopy with extended depth of field. *Opt. Commun.* **268**, 253–260 (2006).
- Thériault, G., Cottet, M., Castonguay, A., McCarthy, N. & De Koninck, Y. Extended two-photon microscopy in live samples with Bessel beams: steadier focus, faster volume scans, and simpler stereoscopic imaging. *Front. Cell. Neurosci.* **8**, 139 (2014).
- Lu, R. *et al.* Video-rate volumetric functional imaging of the brain at synaptic resolution. *Nat. Neurosci.* <http://dx.doi.org/10.1038/nn.4516> (2016).
- Mukamel, E.A., Nimmerjahn, A. & Schnitzer, M.J. Automated analysis of cellular signals from large-scale calcium imaging data. *Neuron* **63**, 747–760 (2009).
- Maruyama, R. *et al.* Detecting cells using non-negative matrix factorization on calcium imaging data. *Neural Netw.* **55**, 11–19 (2014).
- Pnevmatikakis, E. & Paninski, L. Sparse nonnegative deconvolution for compressive calcium imaging: algorithms and phase transitions. *Adv. Neural Inf. Process. Syst.* 1250–1258 (2013).
- McGloin, D. & Dholakia, K. Bessel beams: diffraction in a new light. *Contemp. Phys.* **46**, 15–28 (2005).
- Apthorpe, N.J. *et al.* Automatic neuron detection in calcium imaging data using convolutional networks. Preprint at <http://arxiv.org/abs/1606.07372/> (2016).
- Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Series B Stat. Methodol.* **58**, 267–288 (1996).
- Pnevmatikakis, E. *et al.* A structured matrix factorization framework for large scale calcium imaging data analysis. Preprint at <http://arxiv.org/abs/1409.2903/> (2014).
- Dombeck, D.A., Harvey, C.D., Tian, L., Looger, L.L. & Tank, D.W. Functional imaging of hippocampal place cells at cellular resolution during virtual navigation. *Nat. Neurosci.* **13**, 1433–1440 (2010).
- Yang, W. *et al.* Simultaneous multi-plane imaging of neural circuits. *Neuron* **89**, 269–284 (2016).
- Yaksi, E. & Friedrich, R.W. Reconstruction of firing rate changes across neuronal populations by temporally deconvolved Ca<sup>2+</sup> imaging. *Nat. Methods* **3**, 377–383 (2006).
- Vogelstein, J.T. *et al.* Spike inference from calcium imaging using sequential Monte Carlo methods. *Biophys. J.* **97**, 636–655 (2009).
- Deneux, T. *et al.* Accurate spike estimation from noisy calcium signals for ultrafast three-dimensional imaging of large neuronal populations *in vivo*. *Nat. Commun.* **7**, 12190 (2016).
- Oñativia, J., Schultz, S.R. & Dragotti, P.L. A finite rate of innovation algorithm for fast and accurate spike detection from two-photon calcium imaging. *J. Neural Eng.* **10**, 046017 (2013).
- Grewe, B.F., Langer, D., Kasper, H., Kampa, B.M. & Helmchen, F. High-speed *in vivo* calcium imaging reveals neuronal network activity with near-millisecond precision. *Nat. Methods* **7**, 399–405 (2010).
- Pachitariu, M. *et al.* Extracting regions of interest from biological images with convolutional sparse block coding. *Adv. Neural Inf. Process. Syst.* **24**, 1745–1753 (2013).
- Pati, Y., Rezaifar, R. & Krishnaprasad, P. Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition. *Asilomar Conf. Signals Syst. Comput.* **1**, 40–44 (1993).
- Needell, D. & Tropp, J.A. CoSaMP: iterative signal recovery from incomplete and inaccurate samples. *Appl. Comput. Harmon. Anal.* **26**, 301–321 (2009).
- Donoho, D.L., Tsai, Y., Drori, I. & Starck, J.-L. Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit. *IEEE Trans. Inf. Theory* **58**, 1094–1121 (2012).
- Prevedel, R. *et al.* Fast volumetric calcium imaging across multiple cortical layers using sculpted light. *Nat. Methods* **13**, 1021–1028 (2016).
- Hopt, A. & Neher, E. Highly nonlinear photodamage in two-photon fluorescence microscopy. *Biophys. J.* **80**, 2029–2036 (2001).
- Ji, N., Magee, J.C. & Betzig, E. High-speed, low-photodamage nonlinear imaging using passive pulse splitters. *Nat. Methods* **5**, 197–202 (2008).
- Podgorski, K. & Ranganathan, G.N. Brain heating induced by near infrared lasers during multi-photon microscopy. *J. Neurophysiol.* **116**, 1012–1023 (2016).
- Kim, C.K. *et al.* Prolonged, brain-wide expression of nuclear-localized GCaMP3 for functional circuit mapping. *Front. Neural Circuits* **8**, 138 (2014).
- Low, R.J., Gu, Y. & Tank, D.W. Cellular resolution optical access to brain regions in fissures: imaging medial prefrontal cortex and grid cells in entorhinal cortex. *Proc. Natl. Acad. Sci. USA* **111**, 18739–18744 (2014).
- Cizmár, T. & Dholakia, K. Axial intensity shaping of a Bessel beam. *Proc. SPIE* **7400**, 74001Q (2009).

## ONLINE METHODS

**Microscope design.** The vTwINS microscope was modeled in ZEMAX (Zemax LLC) and custom MATLAB (Mathworks) scripts. The microscope (**Fig. 1c**) was constructed as a modification of a resonant-scanning two-photon microscope. A beam-shaping module to produce the V-shaped PSF for vTwINS was designed to be inserted between the laser and microscope. This strategy was used so that the module could, in principle, be straightforwardly adapted for any existing standard two-photon microscope. The beam-shaping module consisted of three optical paths that could be switched via flip-mount mirrors among (i) a standard high-NA path for standard two-photon imaging, (ii) a vTwINS path using low-NA Gaussian beams, or (iii) a vTwINS path using Bessel beams.

The collimated Gaussian laser beam entering the beam-shaping module had a measured knife-edge width (10–90% power) of 1.3 mm, corresponding to a  $1/e^2$  diameter of 2 mm. The high-NA path consisted of a 2.5× beam expander (AC254-40-B and AC254-100-B, Thorlabs). The Gaussian vTwINS path consisted of a 0.3–1.2× variable telescope (G06-203-525 AC 140/31.5 Linos, LC1120 and AC254-125-B, Thorlabs). When aligning the Gaussian vTwINS path, care was taken to avoid focusing the laser beam directly onto the scanners. The Bessel vTwINS path consisted of an axicon and achromat lens pair (179.2° BK7 Axicon, Altechna and AC254-200-B, Thorlabs) to generate the ring-shaped excitation for the Bessel beam. The specific choice of axicon and achromat lens pair was based on a trade-off between lateral resolution and two-photon excitation efficiency. For the Bessel beams to be correctly formed within the sample, the rear pupil of the objective must be illuminated with well-focused annuli of light. For this reason, the back aperture of the objective is conjugate to the achromatic-lens front focal plane of the axicon–achromat pair. If collimated, parallel beams are used, the two branches of the PSF form an X shape. The PSF V shape was obtained by introducing a slight beam convergence at the objective back aperture created and tuned by a 1× telescope (2× AC254-100-B, Thorlabs). When the vTwINS modalities were used, the beam was split into two parallel beams with a half-wave plate and a calcite beam displacer (AHWP05M-980 and BD27, Thorlabs). The half-wave plate was oriented such that the fluorescence intensities of the two images were equal. The birefringent beam displacer was mounted in a rotation mount and oriented such that the two beams lay in a plane perpendicular to the resonant (fast)-scanning mirror axis of rotation, to ensure that the two images formed from a fluorescent object lay on the same scanned line. A pair of BK7 windows mounted on orthogonal rotation axes was used to adjust and center the lateral position of the beams on the scanners. The beam separation (2.7 mm out of the calcite beam displacer) was further decreased with an 0.8× telescope (AC254-100-B and AC254-80-B, Thorlabs). This specific choice, in combination with the magnification of the microscope (×3.75) and the 12.5-mm focal length of the water-immersion Nikon objective resulted in an angle of 43° between the two branches of the PSF. This choice of angle resulted in an accurate axial localization of the cell bodies (**Supplementary Fig. 2**). When the high-NA path was used for conventional two-photon imaging, the half-wave plate was rotated to zero the power of one of the emerging beams, and the two BK7 windows were oriented to center the remaining beam on the optical axis of the microscope.

A Ti:Sapphire laser (Chameleon Vision II, Coherent) at 920 nm was used for two-photon excitation, and dispersion compensation in the laser was adjusted to maximize the two-photon signal. A Pockels cell (model 350-80 with 302RM driver, Conoptics) was used to modulate the laser intensity, and a half-wave plate plus polarizing beamsplitter cube (Thorlabs) was used to adjust the maximum laser intensity. The two-photon microscope body consisted of a resonant-scanning head (6215/CRS 8 kHz, Cambridge Technologies), a 10-mm  $f$ – $\theta$  scan lens (4401-464-000, Linos) and a 375-mm achromat pair tube lens (2× PAC097, Newport), and an objective lens (N16XLWD-PF, Nikon<sup>41</sup>). The excitation and emission were separated by a short-pass dichroic filter (T680-DCSPXR-UF3 52 mm × 75 mm × 3 mm, Chroma), and the collection optics (ACL7560-A, LC1611-A, ACL25416U-A, Thorlabs) focused the emitted light onto two PMTs (H10770PA-40, Hamamatsu) separated into red and green channels (FF555-Di03-40x54, FF01-720/SP-50, FF02-525/40-32, FF01-593/40-32, Semrock). The PMT signal was amplified with an 80-MHz preamplifier (DHPCA-100, Femto) and digitized with an FPGA (NI PXIe-7961R and NI 5732 DAQ, National Instruments). Scanning and data acquisition were controlled with Scanimage 2015 (Vidrio). The average power during vTwINS data acquisition varied between 150 mW and 200 mW at 920 nm, and the average power during high-NA acquisition was between 50 mW and 70 mW at 920 nm. Images here were typically acquired at 30 Hz with an image size of 512 × 512 pixels with a 90% spatial cutoff, corresponding to an image size of 470 μm × 470 μm (2.8× zoom) or 550 μm × 550 μm (2.4× zoom). Nearly simultaneous calcium imaging through rapid switching between vTwINS excitation and the traditional focused high-NA Gaussian PSF was performed with an alternate optical setup (**Supplementary Fig. 1b**). A galvanometer (6210H, Cambridge Technologies) was used to select between high-NA and vTwINS (38-μm-long FWHM PSF) paths, which were recombined downstream with a (50 μm, 0.88° optical) offset. A modified Scanimage analog control was used to switch between the two paths at every frame (≈17 ms). For each modality, images were acquired at 30 Hz with a 512 × 256 pixel image size. Rapid alternation between two low-NA Gaussian beams was performed with a similar alternate optical setup (**Supplementary Fig. 1c**). Each of the two paths consisted of identical variable low-NA Gaussian beams that were separated by a fixed distance and imaged onto the scanners. The fixed separation distance set the relative angle of the two axially extended beams.

**Transgenic mice.** All experimental procedures were approved by the Princeton University Institutional Animal Care and Use Committee. Transgenic GCaMP6f-expressing mice were produced by crossing Emx1-Cre (B6.129S2-Emx1<sup>tm1(cre)Krl/J</sup>, Jax no. 005628), CaMK2-tTA (B6.Cg-Tg(Camk2a-tTA)1Mmay/DboJ, Jax no. 007004), and TITL-GCaMP6f (Ai93; B6.Cg-Igs7<sup>tm93.1(tetOGCaMP6f)Hze/J</sup>, Jax no. 024103) strains<sup>42</sup>. Male or female transgenic mice heterozygous for all three genes were used at 3–6 months of age for all experiments. No statistical method was used to predetermine sample size. The experiments were not randomized and were not performed with blinding.

**Imaging the mouse visual cortex.** For imaging of the mouse visual cortex, mice underwent surgery under isoflurane anesthesia for implantation of imaging windows and head plates.



A 5-mm-diameter craniotomy was made over one hemisphere of the parietal cortex (centered 2 mm caudal, 1.7 mm lateral to bregma). A custom titanium head plate and optical window (no. 1 thickness, 5-mm-diameter glass coverslip, Warner Instruments) bonded to a steel ring (0.5-mm thickness, 5-mm diameter, SS316 ring, Ziggy's Tubes and Wires) were attached to each mouse's skull with dental cement (Metabond, Parkell). The location of V1 was estimated by using a separate wide-field imaging microscope to record retinotopic responses in fluorescence activity as each mouse viewed horizontally and vertically drifting bars on a 32-inch monitor<sup>43</sup>. Boundaries between the primary and secondary visual areas were defined with an automated algorithm to locate reversals in the retinotopic gradients<sup>44</sup>. Five days after surgery, mice were trained to run on a spherical treadmill (8-inch-diameter Styrofoam ball) surrounded by a 270° toroidal screen<sup>45</sup>. Visual stimuli were generated with the Psychophysics Toolbox<sup>46–48</sup> and displayed on the toroidal screen with a DLP projection system (Mitsubishi HC3000), consisting of ~100 randomly placed and oriented Gabor patches, with visual field size of 5–10°, updated at 4 Hz. To prevent light from the projected display from entering the fluorescence collection system, the region between the base of the objective lens and the head plate was light-proofed with a black rubber tube before imaging. The rubber tube was glued to a silicone ring, and the ring itself was attached to the titanium head plate with silicone elastomer (Body Double, Smooth On). Examples of images from cortical imaging are shown in **Supplementary Figure 7a,c** and **Supplementary Videos 1–4**.

**Imaging the mouse hippocampus.** For imaging the mouse hippocampus, mice under isoflurane anesthesia underwent surgery for implantation of an imaging window and a head plate for head restraint in virtual reality<sup>49</sup>. Optical access to the hippocampus was obtained as previously described<sup>23</sup>. Briefly, an ~3-mm-diameter circular craniotomy over the left hemisphere was performed, centered 1.8 mm lateral to the midline, and 2.0 mm posterior to the bregma. The cortical tissue overlying the hippocampus was aspirated, and a circular metal cannula with a no. 1 coverslip bonded to the bottom was implanted, with a thin layer of Kwik-sil (WPI) placed between the hippocampus and the coverslip. During the surgery, a titanium head plate was attached to the skull with Metabond. After recovery, mice were water restricted for 5 d and then trained to run on a 4-m virtual linear track in a virtual reality setup<sup>50</sup>. Visually distinct towers were placed every 1 m, and 4 µl water rewards were given at 1.6 m and 3.6 m down the track. Mice ran on a 6-inch-diameter Styrofoam cylinder (The Baker's Kitchen) whose position was detected by an angular encoder. Mice were trained for one 60-min session per day and were given 1–1.5 mL of water in total per day (including behavioral training and supplemental water). The virtual-reality projection system was as previously described<sup>44,48</sup> and was controlled with ViRMen<sup>51</sup>. Light-proofing around the objective was performed as described for the experiments in the visual cortex. Examples of images from hippocampal imaging are depicted in **Supplementary Figure 7b** and **Supplementary Videos 5 and 6**.

**Motion correction and preprocessing.** All video sequences were first subjected to a normalized cross-correlation-based motion-correction algorithm. This algorithm, implemented via the template-matching function of OpenCV<sup>52</sup>, found the best horizontal and vertical shifts for each frame to match a fixed template. The

template used was set to the median across frames. Shifts were set to have a maximum allowable value (set to 10 pixels for the V1 data and 15 pixels for the CA1 data). Videos were cropped to remove edge rows and columns with missing data due to shifting. To improve SNR and run time, fivefold temporal running averages and a fivefold spatial binning were applied after motion correction.

For the alternating-beam variation of vTwINS, there were two interleaved videos,  $Y_{right}$  and  $Y_{left}$ . To use SCISM, we preprocessed the data by linearly interpolating each video temporally up to twice the frame rate. The frames at each time step were concatenated side by side, thus creating a vTwINS video of twice the width, in which the minimum distance between pairs,  $d_{min}$ , was the actual distance between beams plus the entire width of the FOV. In the second interleaved video (**Supplementary Figs. 4 and 13d**),  $d_{min} = 15$  pixels, approximating the conditions of a typical vTwINS movie.

**vTwINS orthogonal matching pursuit.** In this section, we describe the mathematical details of the vTwINS SCISM demixing algorithm (**Fig. 3**, **Supplementary Fig. 5** and **Supplementary Video 7**). Let  $Y \in R^{N \times T}$  denote the calcium video sequence,  $X \in R^{N \times K}$  denote the neural spatial components (spatial profiles), and  $S \in R^{T \times K}$  denote the neural temporal activity traces, where  $N$  is the number of pixels in each image,  $T$  is the number of images (or time points), and  $K$  is the number of neurons. Thus, the columns of  $Y$  represent single frames of the video, the columns of  $X$  represent individual spatial profiles, and the columns of  $S$  represent temporal activity traces of single neurons. We model background activity with a set of  $B$  background components  $X_{bg} \in R^{N \times B}$  and denote the (inferred) background temporal activity  $S_{bg} \in R^{T \times B}$ .

Our algorithm is designed to exploit *a priori* knowledge of both the spatial-profile shapes and neural firing statistics. Specifically, the algorithm seeks to factor the full movie matrix,  $Y$ , into the set of spatial profiles,  $X$ , and time traces,  $S$ , such that:

1. The sum of the outer products of spatial profiles and time traces explains the observed data ( $Y \approx XS^T$ ).
2. The time traces,  $S$ , are sparse in time.
3. The spatial profiles are shaped similarly to pairs of neuronal soma (disks or annuli), offset horizontally by a small separation distance. The dark center in each soma is due to the lack of GCaMP6f in the nucleus.
4. Few latent sources (active neurons) relative to the size of the data generate activity in the observed data, thus making the fluorescence movie low rank. This constraint captures the physical density constraints on neuron tissue.

The optimization program that includes all these terms is

$$\{\hat{X}, \hat{S}, \hat{X}_{bg}, \hat{S}_{bg}\} = \arg \min_{X, S, X_{bg}, S_{bg} \geq 0} [\|Y - XS^T - X_{bg}S_{bg}^T\|_F^2 + \lambda_d \|X - D\|_F^2 + \sum_k (\lambda_{gs} \|s_k\|_2 + \lambda_{sp} \|s_k\|_1)] \quad (1)$$

where  $s_k$  is the  $k^{\text{th}}$  column of  $S$ , representing the activity trace of neuron  $k$ ;

$$\|Z\|_F^2 = \sum_{i,j} Z_{i,j}^2 \quad (2)$$

is the squared Frobenius norm;  $D$  is a matrix whose columns represent all possible expected neural spatial-profile shapes;  $\lambda_d$  is

the trade-off parameter for penalizing the deviation of spatial-profile shapes  $X$  from the idealized shapes in  $D$ ;  $\lambda_{gs}$  is the group sparse penalization parameter used to ensure that not all spatial profiles are active; and  $\lambda_{sp}$  is the penalization parameter, which encourages sparsity of the time traces. We set the spatial profiles  $d_k$  to be annuli separated by a depth-dependent distance (Fig. 3a and Supplementary Note 6).

Direct optimization of equation (1) can be inefficient, owing to the problem size and the large search space (number of possible spatial profiles). We thus approximated a solution to equation (1) with a greedy iterative approach wherein spatial profiles are selected sequentially. Our method alternates between finding the best element of  $D$  that approximates  $Y$  given the sparsity constraints (Fig. 3b,c and Supplementary Note 6), updating that profile to the data (Fig. 3d and Supplementary Note 6), and inferring the temporal activity of each spatial profile (Fig. 3e and Supplementary Note 6). The first step sets  $X = D$  and solves for the best single trace to approximate  $Y$  (solving the first and third terms). The shape-refinement step then uses the first two terms with the newly found time trace to allow the spatial-profile  $x_k$  to deviate from its mean  $d_k$ . SCISM is in essence a modification of the orthogonal matching pursuit (OMP) method for greedy sparse signal estimation<sup>31,53</sup>. Our method extends OMP by including an additional temporal sparsity penalty and a shape-refinement step that allows for deviations from the stereotyped neuronal shapes (whereas traditional OMP assumes a fixed dictionary of features).

SCISM was implemented in MATLAB and made use of the TFOCS library<sup>54</sup> to solve the weighted, non-negative LASSO optimization step. Typical analysis ran at a rate of approximately 20 s per profile found and was applied to 10–15 min of  $256 \times 256$  raw imaging data.

**vTwINS and high-NA spatial-profile registration.** High-resolution anatomical  $z$  stacks (median of 200–300 frames per slice at a slice separation of 2.5–4  $\mu\text{m}$ , taken with the high-NA beam path) were obtained for each vTwINS imaging volume to align the vTwINS spatial profiles to anatomical positions. Alignment between the anatomical  $z$  stack and the vTwINS imaging volume was performed in two steps. First, the 3D positions of cells was estimated according to their positions within the vTwINS volume. Second, the estimated 3D positions were offset to the anatomical volume. First, the centroids of each half of the spatial profile were used to calculate the 3D cell position via  $d = 0.5(\Delta - \Delta_{\min})/\tan(\theta)$ , where  $\Delta_{\min}$  is the minimum interbeam distance of the PSF, and  $\theta$  is the beam angle from the axial direction. A correction to the  $x, y$  position was made for any differences in  $\theta$  between the two halves of the vTwINS PSF. Second, a 3D offset between the estimated positions and anatomical  $z$ -stack positions was either automatically or manually calculated. For automatic alignment, the anatomical stack was first deconvolved (Lucy–Richardson) with the high-NA PSF and then convolved with the vTwINS PSF. A 3D cross-correlation was then calculated between the convolution stack and the median vTwINS image, and the peak of the cross-correlation was used as the offset between the vTwINS images and the anatomical  $z$  stack. For manual alignment, highly active cells with similar cell shapes between the vTwINS spatial profiles and high-NA anatomical  $z$  stacks were located manually and used to estimate the offset between the vTwINS images and the anatomical  $z$  stack.

For simultaneous vTwINS and conventional TPM imaging, neural activity was independently extracted from raw images with separate analyses. Neural activity underlying calcium dynamics for conventional TPM was estimated by using the constrained non-negative matrix factorization and deconvolution algorithm (CNMF) to demix contributions from possibly overlapping cells<sup>22,55</sup>. Spatial profiles extracted through CNMF were manually selected for regions that approximated a cell shape (approximately circular, 10–15  $\mu\text{m}$  in diameter). To compare the number of spatial profiles between imaging modalities, spatial profiles from either method were included only if their center position was within 20 pixels (18  $\mu\text{m}$ ) of the  $x$  (fast-scanning) edge of the acquisition region, to prevent bias from clipping half of a single vTwINS profile near the edges of the image.

Spatial profiles and time traces extracted through vTwINS SCISM and CNMF were paired off according to their normalized-time-trace Pearson correlation (Supplementary Fig. 10), with the constraint that the extracted spatial-profile center positions were within 5 pixels (4.5  $\mu\text{m}$ ) in the  $y$  (slow-scanning) direction and 40 pixels (37  $\mu\text{m}$ ) in the  $x$  (fast-scanning) direction. This distance was approximately equal to half the maximum separation distance between vTwINS spatial-profile image pairs, thus not restricting pairing of CNMF spatial profiles to vTwINS spatial profiles with a single blocked beam. Spatial profiles and time traces were paired off until the correlation dropped below a  $5\sigma$  excess of the average correlation between any two time traces. Only high-activity cells with more than one statistically significant transient per minute were included in this analysis<sup>56</sup>. A transient was considered statistically significant if its peak was  $>3\sigma$  above the average noise levels.

**Code availability.** The SCISM source code (written in MATLAB) and documentation on its usage are available on Bitbucket (<https://bitbucket.org/adamshch/scism/>) and as **Supplementary Software**.

**Data availability.** Sample data and the data sets used are available at Open Science Framework under accession code [z6bd3](https://osf.io/z6bd3). Source data files for **Figures 4–6** and **Supplementary Figures 2–4, 6** and **8–14** are available online.

41. Watanabe, K. & Microscope Objective Lens, N.C. Japanese Patent no. 2005-189732 (2005).
42. Madisen, L. *et al.* Transgenic mice for intersectional targeting of neural sensors and effectors with high specificity and performance. *Neuron* **85**, 942–958 (2015).
43. Marshel, J.H., Garrett, M.E., Nauhaus, I. & Callaway, E.M. Functional specialization of seven mouse visual cortical areas. *Neuron* **72**, 1040–1054 (2011).
44. Garrett, M.E., Nauhaus, I., Marshel, J.H. & Callaway, E.M. Topography and areal organization of mouse visual cortex. *J. Neurosci.* **34**, 12587–12600 (2014).
45. Rickgauer, J.P., Deisseroth, K. & Tank, D.W. Simultaneous cellular-resolution optical perturbation and imaging of place cell firing fields. *Nat. Neurosci.* **17**, 1816–1824 (2014).
46. Brainard, D.H. The Psychophysics Toolbox. *Spat. Vis.* **10**, 433–436 (1997).
47. Pelli, D.G. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* **10**, 437–442 (1997).
48. Kleiner, M., Brainard, D. & Pelli, D.G. What's new in Psychtoolbox-3? *Perception* **36**, 1–16 (2007).
49. Harvey, C.D., Collman, F., Dombeck, D.A. & Tank, D.W. Intracellular dynamics of hippocampal place cells during virtual navigation. *Nature* **461**, 941–946 (2009).
50. Domnisoru, C., Kinkhabwala, A.A. & Tank, D.W. Membrane potential dynamics of grid cells. *Nature* **495**, 199–204 (2013).

51. Aronov, D. & Tank, D.W. Engagement of neural circuits underlying 2D spatial navigation in a rodent virtual reality system. *Neuron* **84**, 442–456 (2014).
52. Bradski, G. *Dr. Dobbs J. Softw. Tools Prof. Program*. The OpenCV Library **25**, 120,122–125 (2000).
53. Swirszcz, G., Abe, N. & Lozano, A. Grouped orthogonal matching pursuit for variable selection and prediction. *Adv. Neural Inf. Process. Syst.* **22**, 1150–1158 (2009).
54. Becker, S., Candes, E. & Grant, M. TFOCS: flexible first-order methods for rank minimization. *SIAM Conf. Optim.* (2011).
55. Machado, T.A., Pnevmatikakis, E., Paninski, L., Jessell, T.M. & Miri, A. Primacy of flexor locomotor pattern revealed by ancestral reversion of motor neuron identity. *Cell* **162**, 338–350 (2015).
56. Harvey, C.D., Coen, P. & Tank, D.W. Choice-specific sequences in parietal cortex during a virtual-navigation decision task. *Nature* **484**, 62–68 (2012).