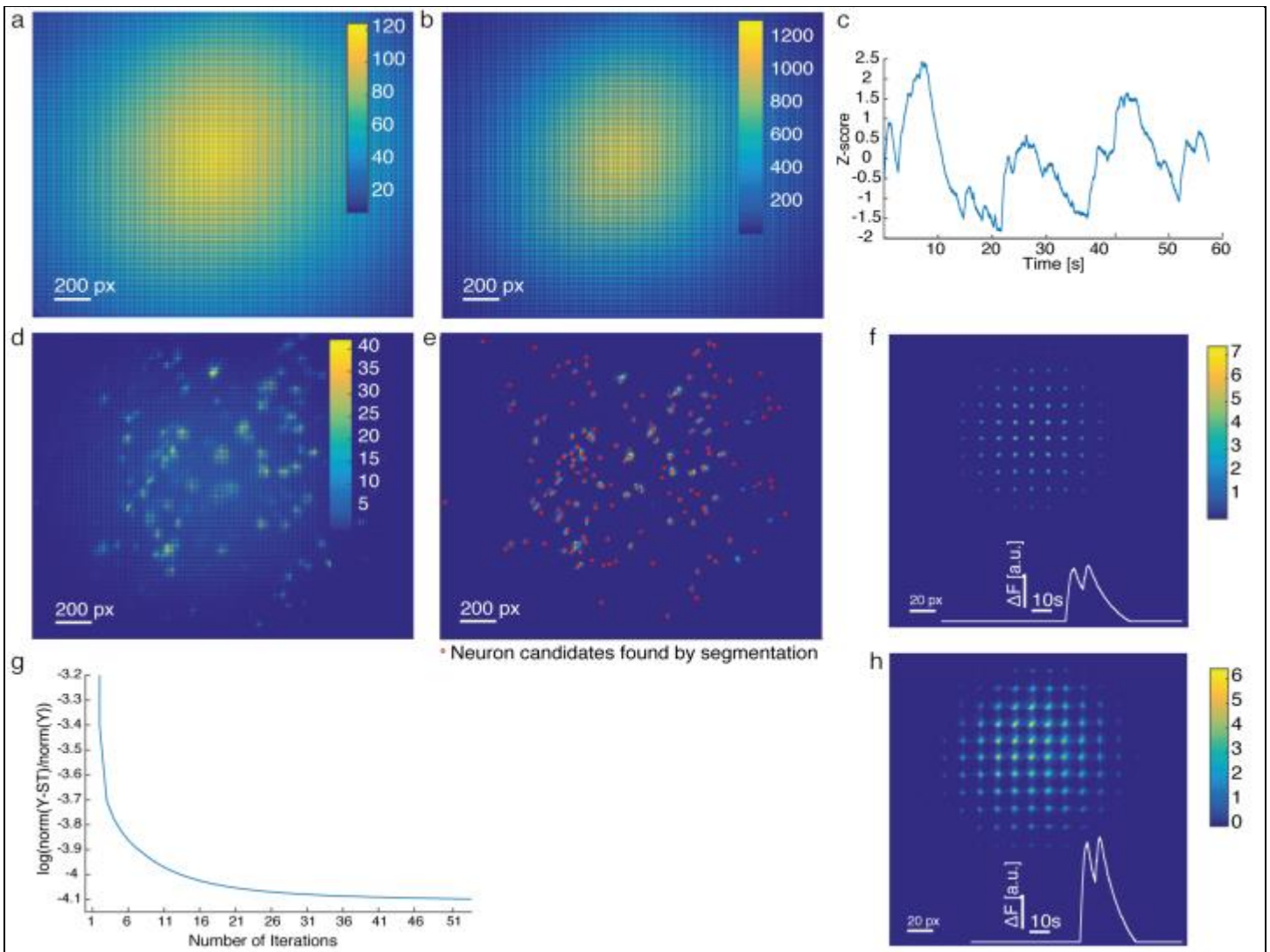**Supplementary Figure 1**

Illustration of the effects of scattering in conventional wide-field and light-field microscopy, illustrating the prior state of the art

(a)   Ground truth: neurons in the mouse cortex.

(b)   Wide field image of ground truth in the presence of scattering.

(c)-(d)   LFM raw image of ground truth without scattering (c) and with scattering (d).

(e)   Zoom into the indicated region in (d), solid white line is intensity profile along dashed white line. Black arrow highlights brightness gradients originating from directionality information retained in scattered light, while white arrow highlights peak due to ballistic light.

(f)   Axial projection of volumetric reconstruction of the LFM image without scattering shown in (c).

(g)   Un-mixed temporal signals extracted from an LFM movie without scattering, from the regions-of-interested indicated as dashed circles in (f).

(h)   Axial projection of volumetric reconstruction of the LFM image with scattering shown in (d).

(i)   Temporal signals extracted from an LFM movie with scattering, from dashed circles in (h). Black rectangles highlight regions with crosstalk from one neuronal signal to the other in the presence of scattering.

Supplementary Figure 2

Seeded source extraction principles and convergence

(a) Standard deviation image of a raw LFM movie from mouse cortex (no background subtraction)

(b) Example of spatial background component extracted using rank-1 matrix factorization of a raw LFM recording from mouse cortex

(c) Example of temporal background component extracted using rank-1 matrix factorization of a raw LFM recording from mouse cortex

(d) Standard deviation image of a background-subtracted LFM movie from mouse cortex

(e) Maximum-intensity projection of LFM reconstruction of the standard deviation image with background subtraction shown in (d). Red circles: Centers of neuron candidates as identified using volume segmentation by local maxima search

(f) Example of initial forward model for a particular neuron candidate. White trace: corresponding initial temporal signal estimate

(g) Convergence plot of alternating spatial and temporal non-negative least squares optimization: Norm of residual vs. iteration count

(h) Example of refined forward model after 20 iterations. White trace: corresponding refined temporal signal estimate
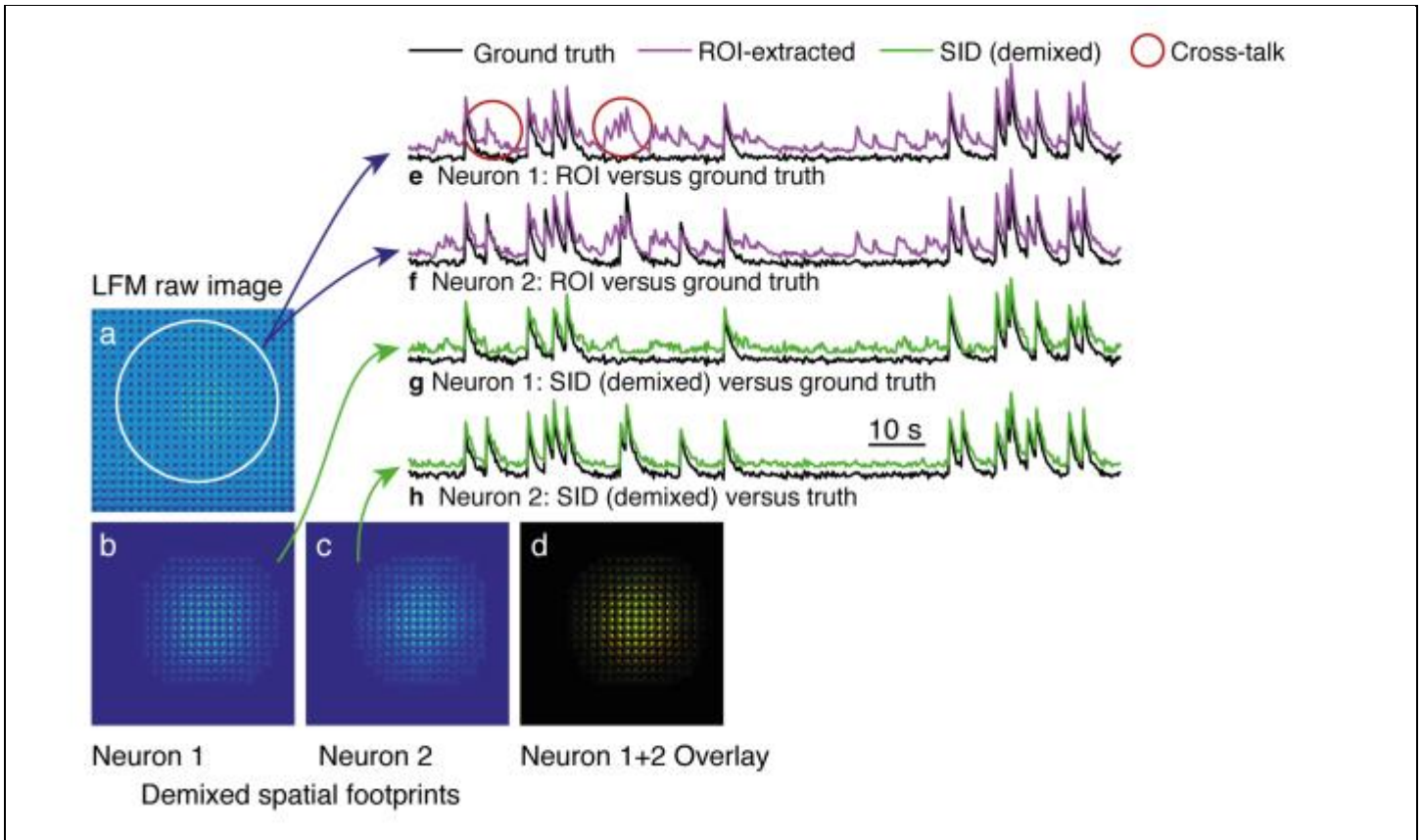
Supplementary Figure 3

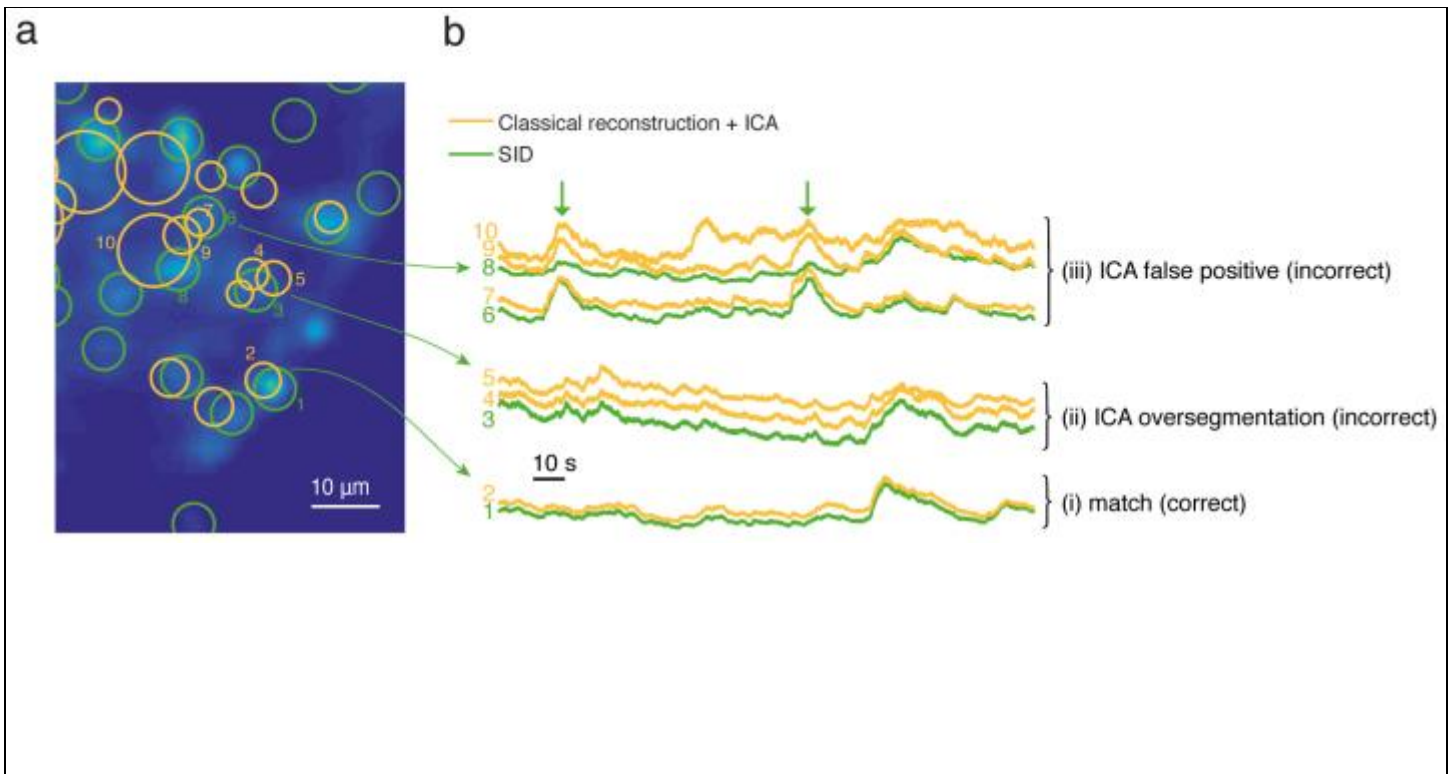Simulation of SID performance versus ground truth: SID-demixing of overlapping neurons with small axial distance

(a) Detail from standard deviation image of a synthetic LFM raw movie generated as described in Suppl. Note 3 for a simulated depth of 450 μm).

(b) + (c) Two demixed spatial components (neuron footprints) detected by SID within the encircled region shown in (a). The detected neurons are spaced 3 μm apart laterally and 8 μm axially in the ground truth.

(d) Composite color overlay of the spatial components (b) and (c), showing considerable overlap (yellow).

(e) - (h) Comparisons of ground truth temporal signals extracted using SID and spatial region-of-interest extraction. Black: Ground truth signals for the two neurons whose LFM images fall into the area encircled in (a). Green: demixed signals corresponding to the spatial components (b) and (c), as indicated by the arrows. Violet: signal extracted by summing over the encircled area in (a). Red circles highlight intervals where time series of neuron 1 (b) and neuron 2 (c) are mixed for a circular region-of-interest extraction (e, f) and demixed for SID (g, h) as evident by comparing the corresponding signals to the ground truth.
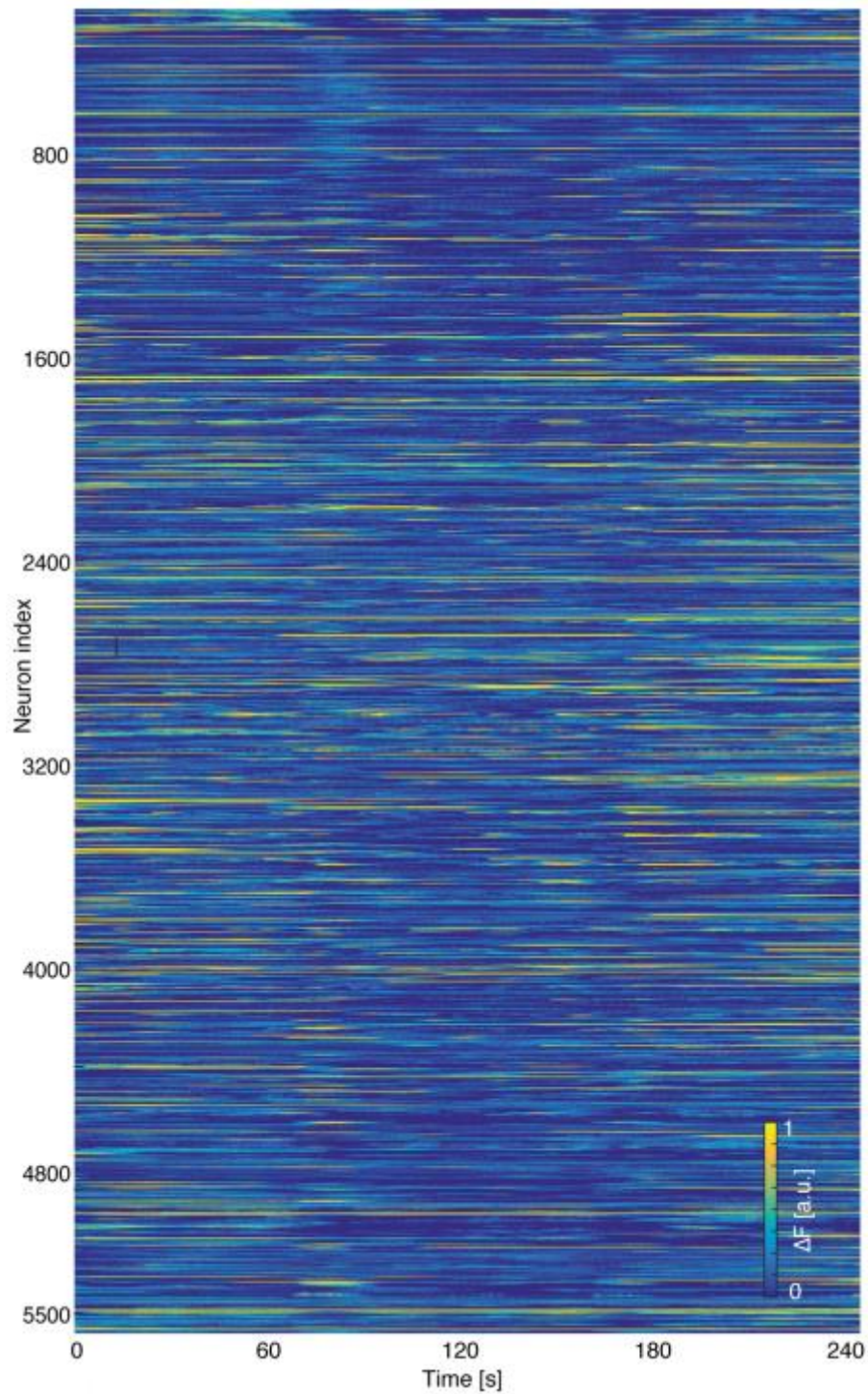
Supplementary Figure 4

Seeded iterative demixing applied to whole-brain Ca²⁺ imaging in larval zebrafish

Heat map of 5505 neuronal signals detected by Seeded Iterative Demixing (SID) in a four minute, 20 Hz LFM recording of spontaneous activity in the larval zebrafish brain. Corresponding neuron locations are shown in Fig. 2 d-e of the main text.

a

b

Classical reconstruction + ICA
SID

(iii) ICA false positive (incorrect)

(ii) ICA oversegmentation (incorrect)

10 s

(i) match (correct)

10 µm

Supplementary Figure 5

Comparison of extraction performance of Seeded Iterative Demixing and an ICA-based method in larval zebrafish

(a) Zoom into a frontal slice of the reconstructed standard deviation image shown in Fig. 2c of the main text, overlaid with spatial component centroids (neuron positions) detected by an ICA-based method (orange circles) and Seeded Iterative Demixing (SID) (green circles). Circle radii are standard deviations of spatial components (orange circles), or a fixed value of 3 µm (green circles).

(b) Example signals detected at the locations encircled in (a). Numbers in (a) and (b) identify corresponding locations and signals. Signals are grouped into three groups (indicated by curly braces) according to their spatial proximity and overlapping spatial components. (i) Example of components found by SID (1) and ICA (2), respectively, that match well both in location and activity. (ii) Example of signals where a single component from Seeded Iterative Demixing (SID) (3) is split up into several overlapping and temporally correlated components by ICA (4, 5). (iii) Examples of components where ICA resulted in signals (9, 10) that appear to be mixtures of signals detected by SID (6, 8), due to ICA detecting "ghost" components in scatter. Vertical arrows indicate time points where peaks from SID-signals (6, 8) appear in ICA-signals (9, 10).
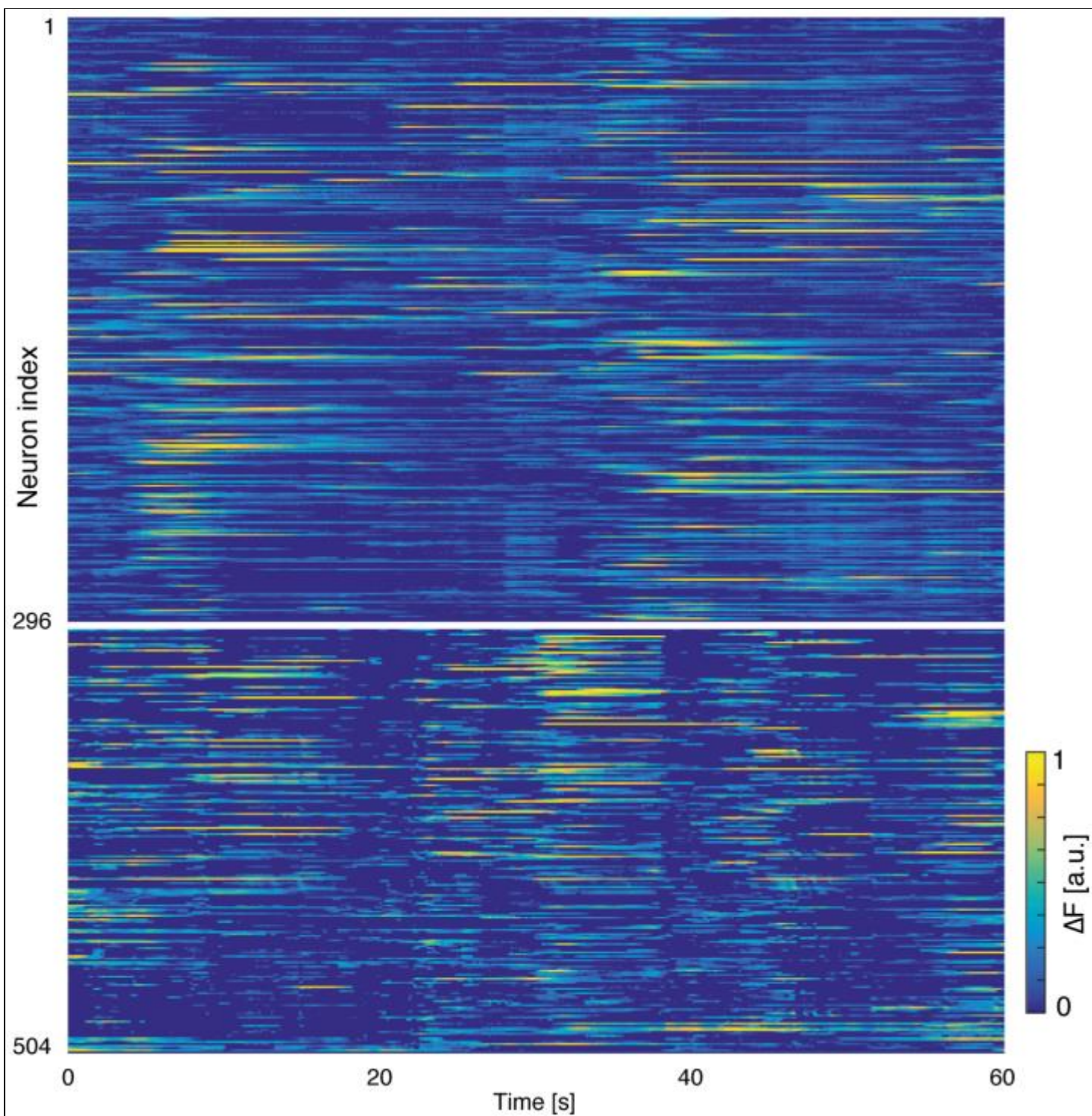
Supplementary Figure 6

Video-rate volumetric Ca^{2+} imaging to 380 μm depth in mouse cortex

Heatmap of neuron activity traces obtained by Seeded Iterative Demixing (SID), corresponding to positions shown in Fig. 3c-d of the main text. Upper panel: 296 active, GCaMP6m-expressing neurons found in 0-170 μm depth range. Lower panel: 208 neurons found in 120-380 μm in a subsequent recording.
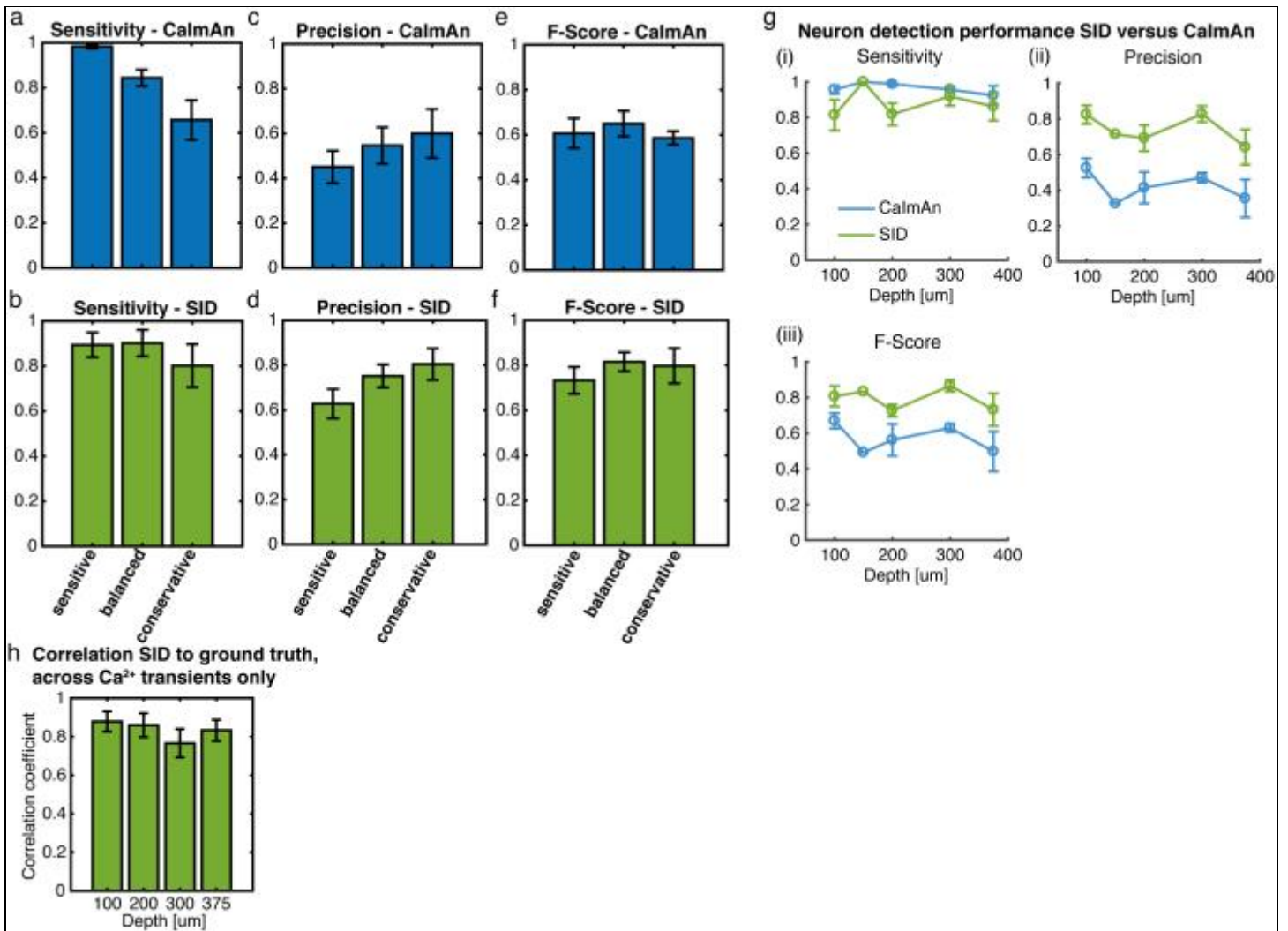
Supplementary Figure 7

Neuron detection scores for different sensitivity settings of SID and CalmAn and quality of SID-extracted Ca^{2+} transients

(a)-(f)	Neuron detection scores "Sensitivity", "Precision" and "F-score" of SID and CalmAn for three different sensitivity settings, as described in Suppl. Note 5
(g)	Neuron detection scores versus depth as achieved by SID (green traces), in comparison to scores achieved by the analysis package *CalmAn* applied to the 2PM data (blue traces), both evaluated with respect to a ground truth. (i) Sensitivity score (ratio of number of detected to actual neurons), (ii) Precision score (ratio of number of true positives to sum of true and false positives), (iii) F-Score (harmonic mean of sensitivity and precision). n = 4.
(h)	Correlation of SID-extracted signals to ground truth (see Suppl. Note 5), computed across peaks only, for four different tissue depths
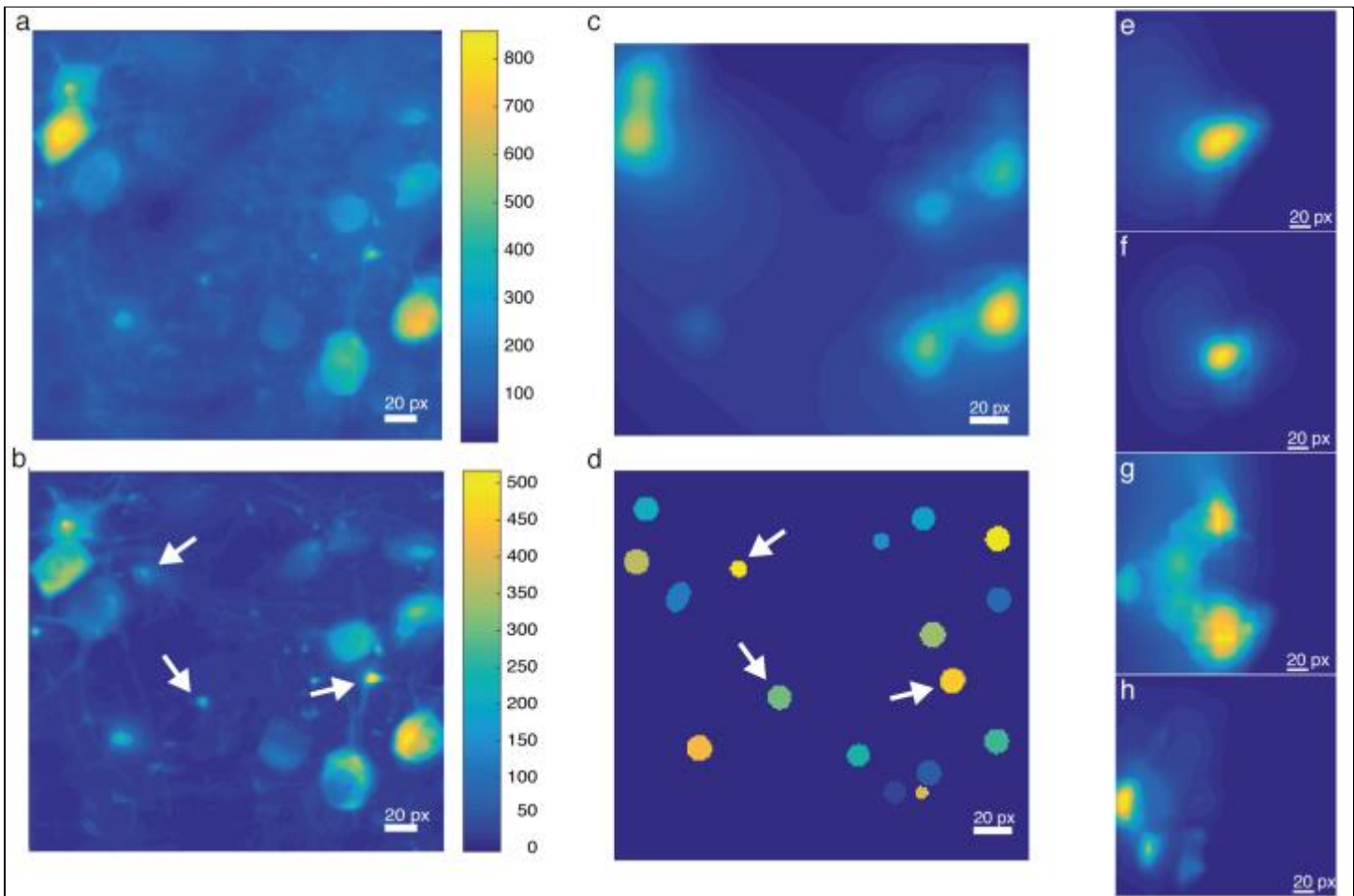
Supplementary Figure 8

Background suppression and neuropil rejection

(a)   Mean image of 2PM movie of a single plane in mouse cortex, same data as in Fig. 5c

(b)   Standard deviation image of 2PM movie of a single plane in mouse cortex, same data as in Fig. 5c. White arrows highlight neurites that are identified by a local maximum search and segmentation of the LFM standard deviation image (c, d), in addition to active somata.

(c)   Maximum-intensity projection of reconstructed standard deviation image of LFM movie from mouse cortex, recorded by sending part of the 2P-excited emission light to an LFM camera in a hybrid 2PM-LFM setup. Corresponding 2PM signal shown in (a) and (b).

(d)   Result of volume segmentation of LFM-reconstructed standard deviation image shown in (c). Comparison with (b) shows that in addition to the active somata, some large and active neurites (corresponding pairs highlighted with white arrows in (b) and (d)) are identified by the local maximum search and segmentation algorithm and thus can be subtracted.

(e) + (f):   Examples of axial maximum intensity projections of LFM-reconstructed footprints classified as neurons

(g) + (h):   Examples of axial maximum intensity projections of LFM-reconstructed footprints classified as neuropil
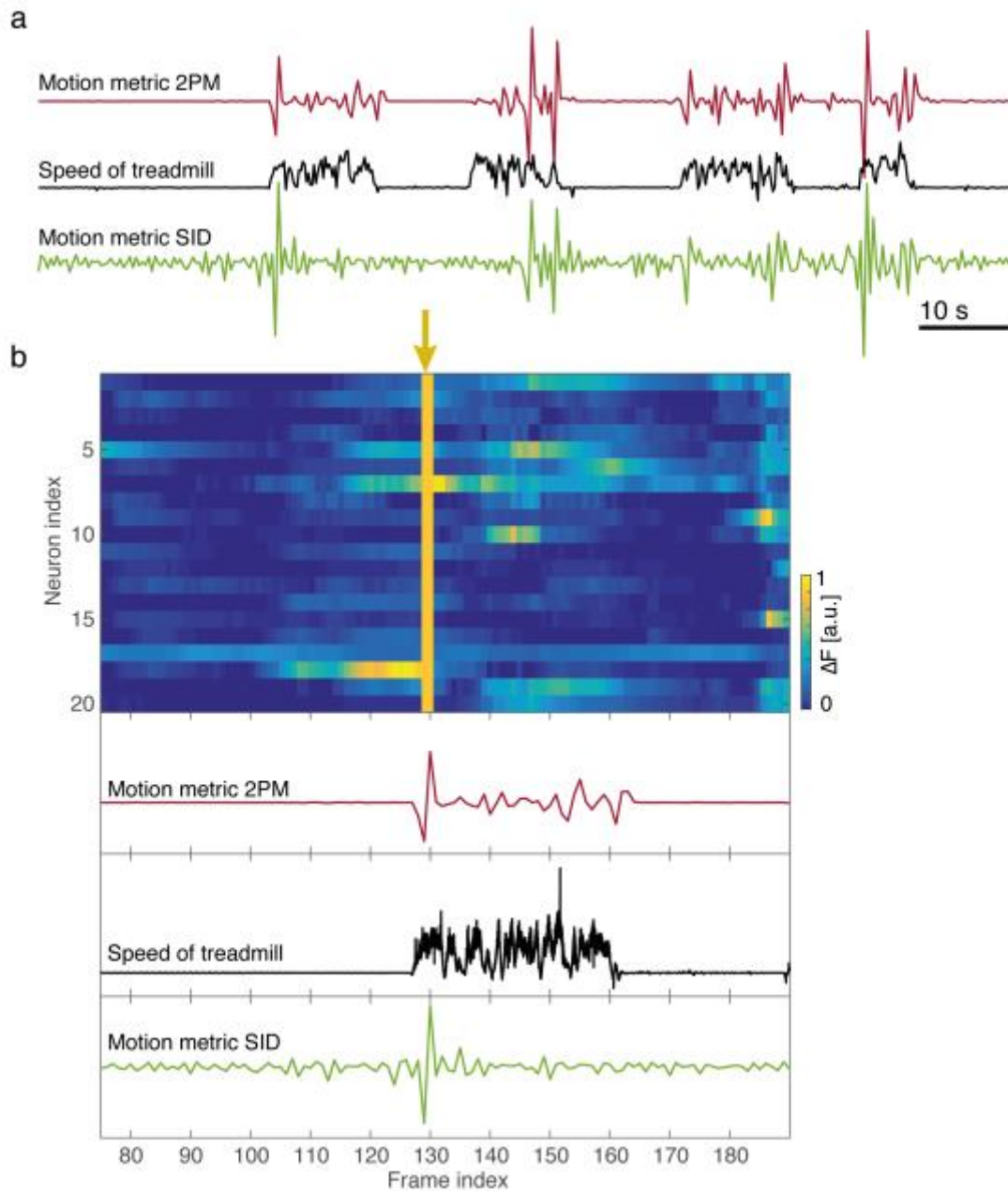
Supplementary Figure 9

Motion detection and motion frame exclusion in SID raw data

(a) Motion metric based on image autocorrelation as described in Suppl. Note 7, evaluated for a 2PM recording (red trace) and a simultaneously acquired SID recording (green trace) from mouse cortex (PPC) at 200 μm depth. Black trace indicates rotational speed [a.u.] of treadmill disk due to walking behavior of a head-fixed mouse, as measured by a high-speed optical computer mouse in proximity to the treadmill disk

(b) Top panel: Heatmap of SID-extracted neuronal activity traces from simultaneous 2P-SID recording (frame rate 3 Hz) from mouse cortex (PPC) at 200 μm depth. Bottom panels: Motion metrics and treadmill tracking data as in (a). Column marked in yellow in top panel indicates frame excluded from further analysis due to high motion metric value in corresponding time bin, as described in Suppl Note 7
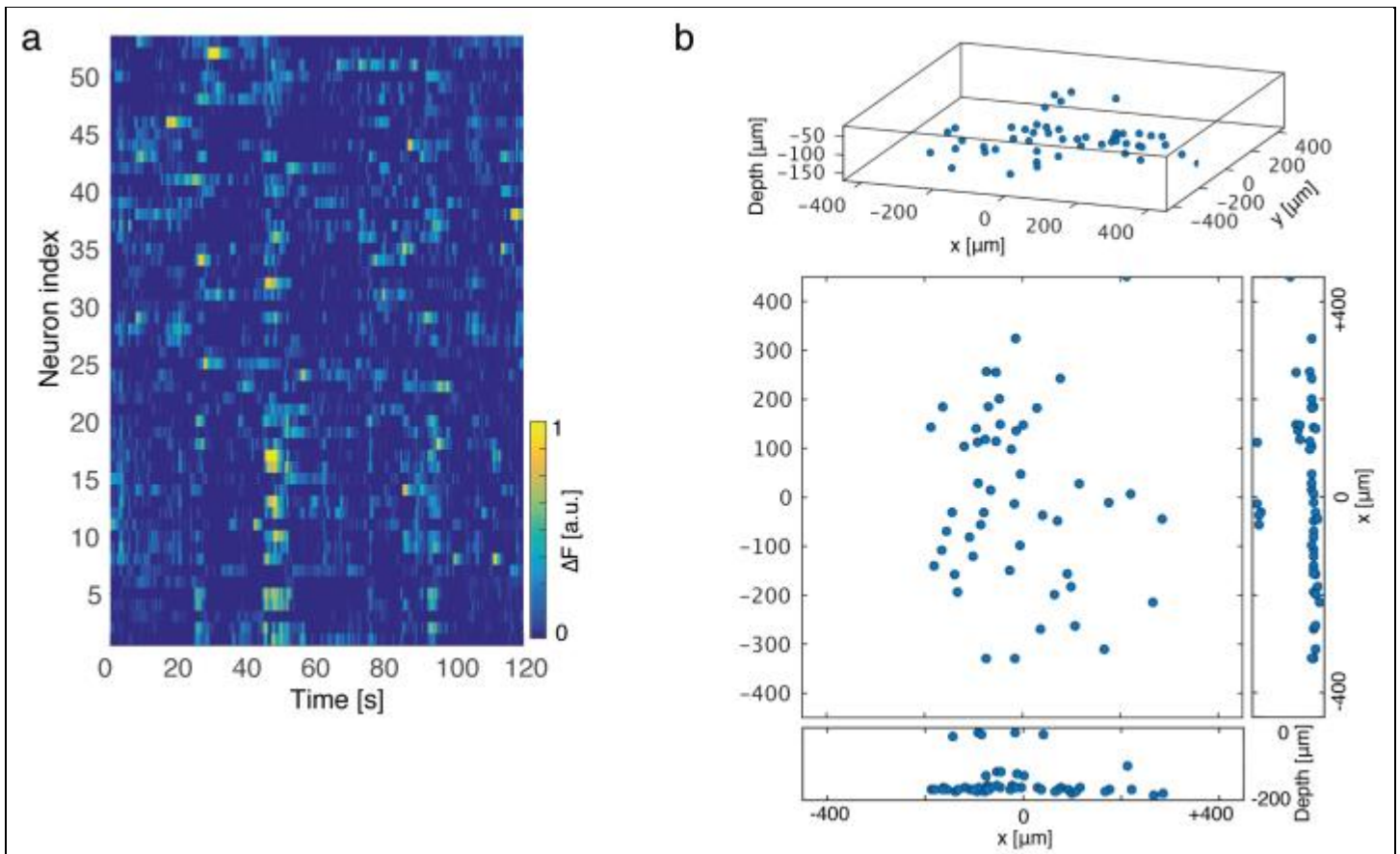
Supplementary Figure 10

SID-extracted positions and signals of jRGECO-labelled neurons in mouse hippocampus CA1

(a)   Heatmap of SID-extracted neuronal activity traces for 54 neurons recorded from mouse hippocampus CA1 at a frame rate of 5 Hz. Cranial window surgery and imaging as described in the main text and Methods section. (The recordings were performed under suboptimal conditions in which the expression of jRGECO was not fully developed and the transfection protocol sub-optimal)

(b)   SID-extracted 3D neuron positions corresponding to activity traces in (a). Top panel: isometric view. Bottom panels: Top, side and front views. Depth is given from bottom surface of cranial window placed onto corpus callosum as shown in Fig. 4a of the main text

Supplementary Notes

Video rate volumetric calcium imaging across multiple cortical layers using Seeded Iterative Demixing (SID) light field microscopy

Tobias Nöbauer[1,*], Oliver Skocek[1,*], Alejandro J. Pernía-Andrade[2], Lukas Weilguny[2], Francisca Martinez Traub[1], Maxim I. Molodtsov[2] & Alipasha Vaziri[1,2,‡]

[1] Laboratory of Neurotechnology and Biophysics, The Rockefeller University, New York, NY, USA
[2] Research Institute of Molecular Pathology, Vienna, Austria
[*] These authors contributed equally to this work
[‡] Correspondence should be addressed to A.V. (vaziri@rockefeller.edu)

SUPPLEMENTARY NOTE 1: SEEDED ITERATIVE DEMIXING (SID) ALGORITHM IMPLEMENTATION DETAILS

Background rejection

Deep tissue LFM movies contain strong global background fluorescence which has to be subtracted before computing a standard deviation image and before any further steps. This background is mostly due to fluorescence originating from above and below the depth range captured by the numerically simulated PSF that is used for reconstruction. We extracted this background by applying a rank-1-matrix-factorization to the LFM raw data. The spatial and temporal components obtained from rank-1-matrix-factorization are added to the neuron candidates in the spatial and temporal update steps – as described in the Methods section – as an additional row and column of the S and T matrices, respectively. The background estimates are therefore refined during these optimization steps, and activity may be re-allocated from neurons to the background, and vice versa. In the temporal update step, this corresponds to an inherent background subtraction, while in the spatial update step, the shape of the background is refined.

Without background subtraction, the standard deviation image of an LFM movie is dominated by temporal variations in the background (Suppl. Fig. 1a). We found a one-dimensional approximation of the background sufficient to obtain the ballistic components of the neuron footprints. Suppl. Fig. 1b illustrates the spatial component (background image) of the rank-1 factorization, and Suppl. Fig. 1c the corresponding temporal component (background signal). We compared the standard deviation image without and with background subtraction, respectively (Suppl. Fig. 1a, 1d). It is evident that removing the background reveals LFM footprints of localized sources.

Reconstruction with sparsity, segmentation

The standard deviation images were reconstructed (de-convolved with numerically simulated PSF) using a modification of a Richardson-Lucy-type algorithm known as ISRA[1], which yields non-negative components. Classical LFM reconstruction based on Richardson-Lucy deconvolution with a ballistic PSF[2,3] is prone to blocky artefacts near the native focal plane of the microscope where the optical spatial sampling density is strongly reduced[2]. These artefacts are detrimental to the success of the subsequent segmentation procedure. When necessary, we therefore modified ISRA with a sparsity constraint. The update step for volume estimate x is:

$$x_{n+1} = x_n \, \frac{P^T y}{P^T P y + \lambda \, 1_{\dim(x)}},$$

where $1_{\dim(x)}$ is a vector of ones with the same dimension as x, and P is the PSF. The parameter λ governs the weight of the sparsity-encouraging term. We used $\lambda > 0$ for the zebrafish recordings shown in Fig. 2. For deep mouse recordings, we set $\lambda = 0$ for performance reasons and instead discarded neuron candidates detected in the artefact region. Before reconstruction, standard deviation images were thresholded to exclude residual background activity.

Segmentation

In order to suppress spatial frequencies not compatible with neuron shapes, a bandpass filter was applied to the reconstructed standard deviation volume, followed by thresholding the result to exclude background. Then, a local maximum search algorithm was applied, Suppl. Fig. 1e illustrates the result: Detected regions in a reconstructed standard deviation image are labelled with red dots. The segmentation threshold is chosen to robustly reject noise and artefacts.

Non-negative matrix factorization

The algorithm proceeds as described in the Methods section of the main text, by alternating temporal and spatial update steps. Suppl. Fig. 1f shows an example of the initial estimates of spatial and temporal components of a neuron. Suppl. Fig. 1h shows the components for the same neuron after two update iterations: While the initial spatial estimate only includes the ballistic footprint, the updated estimate increasingly incorporates the scattered light around it. The corresponding temporal components become more pronounced and – crucially – increasingly de-mixed from overlapping signals (Fig. 1c and 5a).

Convergence

Both the spatial and temporal optimization steps as defined in the Methods section are convex problems and therefore each converge to a global optimum. The combined problem is bi-convex and a variant of what is known as an alternate convex search[4] in the literature, which is a frequently used algorithm for this class of problem. The alternate convex search algorithm optimizes a bi-convex target function by splitting the problem into its convex sub-problems, initializes the solution with a guess, and iteratively solves one of the two sub-problems, while keeping the other variable fixed at the optimum of the previously solved sub-problem (or the initial guess), and then alternating the sub-problems until a stopping criterion is reached. It has been shown[4] that the iteration sequence pursued by the alternate convex search algorithm has at

least one accumulation point, and that if each accumulation point has a unique solution for each of the sub-problems, then the difference between consecutive iterations converges to zero. The value of the target function is the same at each accumulation point, and reaches a partial optimum (i.e., an optimum in each of the convex variables). In a strict sense, the global optimality of the solution is not guaranteed. However, alternate convex search is routinely applied to bi-convex optimization problems, for instance in the context of Ca^{2+} imaging for spatio-temporal demixing of 2PM data[5], with good success.

For both the spatial and temporal update steps, we use the ISRA algorithm without a sparsity constraint. We found it to parallelize efficiently across multiple CPU-cores as well as thousands of GPU-cores, allowing for quick solution of large problems (thousands of pixels times thousands of time steps within approx. 1 GPU-second per neuron). In Suppl. Fig. 1g, the normalized residual of the combined spatio-temporal updates is plotted on a logarithmic axis versus the number of iterations. We routinely observe fast convergence and abort the algorithm after approx. 10 iterations, when the residual has been reduced by four orders of magnitude. At this point, no spatial or temporal structure is evident in the residual data anymore.

SUPPLEMENTARY NOTE 2: SYNTHETIC DATASET GENERATION

The synthetic dataset used for Fig. 1c of the main text was generated as follows, using literature values for the parameters[7–9]: 40 neurons (spheres of 8 μm diameter) were randomly placed in a volume of $70 \times 70 \times 200$ μm, maintaining a minimum distance of one neuron diameter, and surrounded by a margin of 25 μm on each side to avoid border artefacts. The simulated neuron density was chosen to be 40,000 per cubic millimeter. This is lower by a factor of approx. two than the average density reported for mouse cortex[10], to account for the fact that not all neurons are active during a given recording. The volume size was chosen large enough to span most of the LFM axial range, and for scattered neuron images originating from distant sides of the volume to be non-overlapping on the simulated LFM sensor, while keeping computational effort within the capacity of a 20-CPU-core, quad-GPU workstation. Poissonian spike trains of action potentials were randomly generated (mean firing rate 0.5 Hz, 1000 time steps at a 5 Hz sampling rate), linearly mixed to introduce some correlation among them (mixing matrix chosen to result in an exponential distribution of variances explained by principal components), and convolved with an exponentially decaying GECI response kernel (mean decay time constant 1.2 s). Gaussian noise was added to the resulting traces to emulate a GECI signal-to-noise ratio (SNR) of 25.

The randomly placed neurons and the simulated GECI activity traces were then combined to generate a time series of volumes. To account for fluctuations of the background fluorescence due to neuropil and detection noise, a noisy background was added throughout the synthetic volumes (SNR 25), as well as to the final simulated sensor image. To obtain simulated sensor data in the absence of scattering, the synthetic volumes were convolved with a numerically simulated, ballistic LFM PSF (corresponding to a 16x 0.8NA water dipping objective). To obtain an approximation of the scattered sensor data, the synthetic volumes were convolved with a simulated scattered PSF obtained from a Monte-Carlo approach (see next subsection) for a

scattering length of 100 μm, a depth 400 μm, and a Henyey-Greenstein anisotropy parameter 0.9, in accordance with literature values[7,8].

Monte-Carlo simulation of scattered PSF

To generate the scattered PSFs, we followed a Monte-Carlo approach inspired by Ref. [11]: 100000 virtual rays were launched from a point source on the optical axis and propagated by sampling the distances between scattering events (free paths) from an exponential distribution and scattering angles from a Henyey-Greenstein distribution. For each scattering event, a "virtual" source was placed at the apparent origin of the scattered ray and at a depth corresponding to the free path before the scattering event. The resulting volume of virtual sources was projected forward to the sensor by convolving with the ballistic PSF. This was repeated for every lateral and axial displacement necessary to fully capture the spatially varying, but periodic structure of the LFM PSF.

SUPPLEMENTARY NOTE 3: STATISTICAL ANALYSIS OF SID-EXTRACTED NEURONAL SIGNALS

To obtain the extraction quality characterizations presented in Fig. 6 of the main text and Suppl. Fig. 4, we recorded a set of single-plane, simultaneous 2PM-SID movies at a series of depths from the posterior parietal cortex of awake, head-fixed mice (100–375 μm, total n = 18 recordings, 4 animals). In what follows we discuss the analysis steps involved in more detail and discuss the optimal tuning of extraction parameters for robust extraction performance.

Signal extraction and tuning of detection characteristics

The recently published constrained matrix factorization algorithm for Ca^{2+} signal extraction[5] implemented in the *CaImAn* analysis package was used to analyze the 2PM recordings, exactly implemented in the demo script[6] that comes with the package, adapting the neuron size and approximate number of active neurons to values suitable for our data. After running an initialization subroutine and the core constrained matrix factorization, the script performs post-selection of ROIs based on spatial shape and size. We found that the overall sensitivity and precision of the algorithm depends mostly on the thresholds for required convexity and size of neurons, as well as the approximate number of active neurons chosen initially. We determined three sets of parameter values for our data that result in three estimation qualities: A "sensitive" estimate (avoid missing neurons while accepting a greater risk of detecting false positives), a "conservative" estimate (avoid false positives while taking greater risk of missing actual neurons), and a "balanced" setting that aims for the optimal trade-off between sensitivity and precision.

The light-field raw data was processed as described in the main text and Methods section, as well as Suppl. Note 1. After background-subtraction, the motion metric described in Suppl. Note 5 was calculated, and motion-affected frames excluded from further processing. The sensitivity and precision values of SID are tuned by varying two parameters that estimate the noise floor and the background level, respectively, of the data and manually inspecting the output of the segmentation step. Sensitivity can be increased at the expense of precision by the lowering noise floor and background

estimates, and vice versa. Again, three different sets of parameters were chosen that resulted in conservative, balanced and sensitive signal extraction qualities. SID was run with the "balanced" setting on all datasets, and in addition with the "conservative" and "hypersensitive" settings on the recordings from one animal.

Compilation of ground truth and categorization of detections

We manually inspected the output of the sensitive CaImAn runs and categorized the detections contained therein as true or false positives by assessing the shape of the detected object, and whether a single object was segmented into several ROIs. Any neurons that were not picked up were added manually and categorized as false negatives. Together, the true positive CaImAn detections and manually added neurons (positions and signals) in the 2PM recordings constitute what we regarded as the ground truth for all further analyses.

In a second manual step, we assessed all SID runs of the "sensitive" quality setting by comparing SID-detected locations to the ground truth locations, identifying the matching pairs, and adding any missing neurons, marking them as false negatives. The categorizations as true/false positives/negatives of all other CaImAn and SID results (i.e., the "balanced" and "conservative" extraction qualities) were inferred by automatic comparison to the locations and signals that were categorized manually based on the "sensitive" extraction output, followed my manual inspection and verification.

Neuron detection scores

To describe the neuron detection performance of the CaImAn and SID, we computed three standard quality scores commonly used in the context of classification/detection models: The score known as *recall* or *sensitivity* (ratio of true neurons to detected neurons); the *precision* (ratio of true positives to total detections, i.e. to the sum of true and false positives); and the *F-score*, which is defined as the harmonic mean of precision and recall (multiplied by two to scale its value to the (0,1) range). The F-score is one when both sensitivity and precision are equal to one, i.e. all true neurons were detected correctly, and no false positives detections appeared.

In Suppl. Fig. 4a-f, we plot these three scores for both SID and CaImAn, and the three extraction quality settings: While the "sensitive" quality setting maximizes the sensitivity scores in both SID and CaImAn, the "conservative" setting results in a maximal precision scores. The F-scores are optimized for the "balanced" setting. This result verifies that the parameter sets were chosen in an appropriate way, and we determine the "balanced" SID setting to be the default setting in the SID implementation published as Supplementary Code with this article.

Correlation analysis of SID-extracted neuronal signals

For the signal quality assessments presented in Fig. 6b-c of the main text, we computed the zero-lag correlation coefficients of the true positive SID signals and their respective counterparts in the ground truth, including their entire duration. The values given in Fig. 6b-c therefore contain information both about whether any peaks in the extracted signals match with the ground truth peaks (true/false positive GECI transient detections), and on whether their absence in the extracted signal is correct (true/false negative transient detections). For comparison, we also calculated the

correlation of the SID signals to ground truth across peaks only. A histogram of the resulting peak-gated signal correlations versus depth is shown in Suppl. Fig. 4g: In comparison with the ungated data shown in Fig. 6b-i of the main text, we observe no significant differences. This is an indication that any mismatches in the extracted signals compared to ground truth are not strongly biased towards false negative or false positive peaks, and that the ungated correlation values used throughout Fig. 6b-c are a good measure of signal extraction quality.

SUPPLEMENTARY NOTE 4: NEUROPIL REJECTION

Generally, it can be desirable to decontaminate the neuronal signals from that of nearby neurites as well as from any background signals (neuropil). In our approach, diffuse fluorescence from neuropil and very small neurites are rejected to a large degree due to background subtraction and the use of a standard deviation image as the starting point for segmentation but also the remainder of the algorithm. This is illustrated in Suppl. Fig. 2 using a planar movie from mouse cortex recorded simultaneously in LFM and 2PM. While the signal-to-background ratio is as low as ~2 in the mean image of a 2PM planar movie recorded depth 200 μm (Suppl. Fig. 2a), it is as high as ~20 in the standard deviation image of the same movie (Suppl. Fig. 2b): In the latter, diffuse background is strongly suppressed compared to the active cell bodies and larger neurites. The high-intensity regions of the 2PM standard deviation image – which clearly are somata – also stand out in the corresponding reconstructed standard deviation image of the LFM recording (Suppl. Fig. 2c), and reliably get identified by a local maximum search algorithm followed by a segmentation (Suppl. Fig. 2d). This algorithm primarily picks out the active somata, but also some of the larger and very active neurites (highlighted using arrows in Suppl. Fig. 2b and 2d). These larger neurites are processed further, and their spatial and temporal components are optimized iteratively as described above. After the optimization, we can reconstruct the optimized spatial components to more closely examine their shape: While the cell bodies are compact, larger and spherically shaped (Suppl. Fig. 2e and 2f), neurites often extend over a larger region, both due to their morphology and since nearby neurites are often merged into the same spatial component due to their correlated activity, and have less regular shapes (Suppl. Fig. 2g and 2h). These differences are used for manual or automatized post-selection processing whereby the signals from neurites can be identified and subtracted out from that of neuronal cell bodies.

SUPPLEMENTARY NOTE 5: MOTION DETECTION AND CORRECTION

During imaging sessions, mice were head-fixed using a customized mount complemented with a head bar holder and a mouse body stabilizer (body jacket) and could run freely on a disc (200 mm diameter), as described in more detail elsewhere[12]. This considerably reduced animal-induced motion of the brain during imaging. To detect any residual motion in the raw SID/LFM raw data prior to further processing, we developed a simple motion detection metric based on image autocorrelation which is computed as follows: First, the raw data is background-subtracted by rank-1 non-negative matrix factorization of the time series of SID/LFM camera frames. Next, we compute the difference frames between all background-subtracted frames, and

compute the autocorrelation images of the difference frames. In the difference frames, translation of a source within the FOV manifests itself as negative values at pixels illuminated from the previous source position, and positive values at pixels illuminated by the new source position. Hence, the values of these two sets of pixels will be anti-correlated, resulting in a negative peak in the autocorrelation image, at a spatial "lag" (distance) corresponding to the extent of the motion effect. We extract the minima of each autocorrelation image (normalized to the maximum of the autocorrelation image), and take the time derivative of this series of minima to obtain a clear metric for motion in the LFM raw frames. This metric is what is plotted in Suppl. Fig. 5a for data from a simultaneous 2PM+SID recording: The motion metric computed from the of the 2PM and LFM/SID raw data are in good agreement, and the peaks in both metrics correlate with the onset of animal motion as recorded by tracking the movement of the running disc with a high-resolution optical computer mouse (middle trace).

In SID/LFM, the point-spread function of the system is engineered to vary spatially (in order to provide axial resolution), so a translation of a source does not result in a mere translation of the image on the sensor as in classical wide-field imaging, but a more intricate transformation. However, we found that simply taking the minima of the difference frame autocorrelation images still picks up motion well.

Pixels affected by motion would exhibit high standard deviation along time that does not originate from neuronal activity, and would thus negatively affect the precision of SID demixing and segmentation. We therefore exclude frames with a motion metric value above a threshold prior to computing the standard deviation image (step ii in Fig. 1b of the main text). An example of an excluded frame together with motion metrics and running disk tracking data is shown in Suppl. Fig. 5b.

Within the scope of this work we did not attempt to recover neural activity from the motion-affected frames. Since LFM/SID captures the full recording volumes in an unbiased way, we expect it to be possible to recover neuron activity information by registering the SID-detected neuron footprints of the unaffected frames to the transformed footprints in the motion-affected frames and extract the source brightness. As mentioned above, the translation of a source (neuron) in LFM/SID results in transformation of its LFM image that is not a simple translation, due to the spatially varying point-spread function in LFM. However, since the point-spread function is known, it is possible map source positions to images and iteratively find the transformation of source positions that best explains the image observed during motion frames. This procedure can be based on a standard optimizer for image registration, with the additional step of mapping position estimates to LFM images by convolving with the LFM point-spread function.

SUPPLEMENTARY REFERENCES

1. Daube-Witherspoon, M. E. & Muehllehner, G. An Iterative Image Space Reconstruction Algorithm Suitable for Volume ECT. *IEEE Trans. Med. Imaging* **5,** 61–66 (1986).

2. Cohen, N. *et al.* Enhancing the performance of the light field microscope using wavefront coding. *Opt. Express* **22,** 24817–24839 (2014).

3. Prevedel, R. *et al.* Simultaneous whole-animal 3D imaging of neuronal activity using light-field microscopy. *Nat. Methods* **11,** 727–730 (2014).

4. Gorski, J., Pfeuffer, F. & Klamroth, K. Biconvex sets and optimization with biconvex functions: a survey and extensions. *Math. Methods Oper. Res.* **66,** 373–407 (2007).

5. Pnevmatikakis, E. A. *et al.* Simultaneous denoising, deconvolution, and demixing of calcium imaging data. *Neuron* (2016).

6. CaImAn package, demo script. *GitHub* Available at: https://github.com/simonsfoundation/CaImAn. (Accessed: 10th April 2017)

7. Helmchen, F. & Denk, W. Deep tissue two-photon microscopy. *Nat. Methods* **2,** 932–940 (2005).

8. Jacques, S. L. Optical properties of biological tissues: a review. *Phys. Med. Biol.* **58,** R37 (2013).

9. Chen, T.-W. *et al.* Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499,** 295–300 (2013).

10. Schüz, A. & Palm, G. Density of neurons and synapses in the cerebral cortex of the mouse. *J. Comp. Neurol.* **286,** 442–455 (1989).

11. Nairat, M. & Voelz, D. Approach for incorporating aerosol scattering in wave optics propagation simulation. in *2013 IEEE Aerospace Conference* 1–5 (2013). doi:10.1109/AERO.2013.6497321

12. Prevedel, R. *et al.* Fast volumetric calcium imaging across multiple cortical layers using sculpted light. *Nat. Methods* **13,** 1021–1028 (2016).