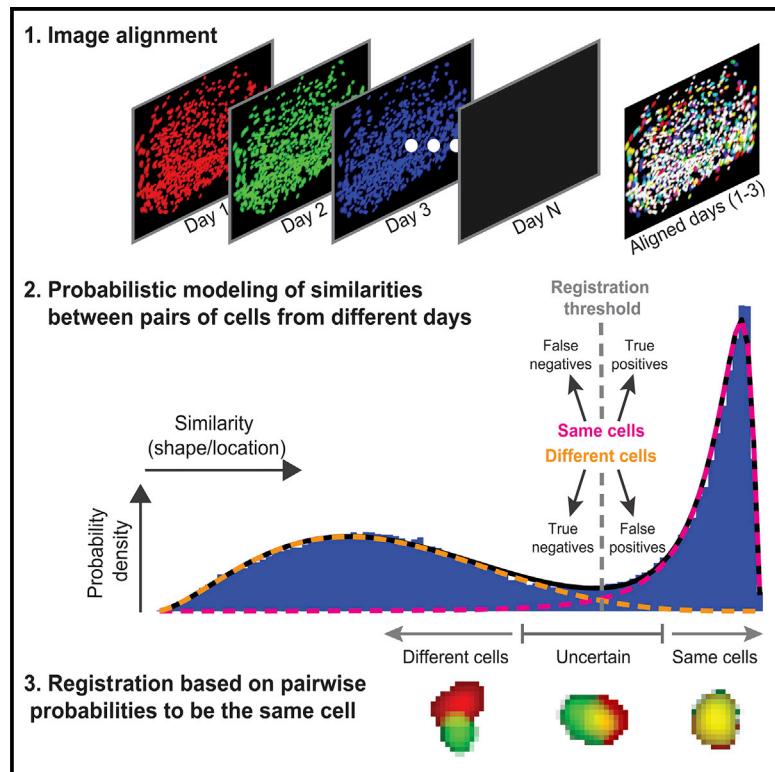


**Tracking the Same Neurons across Multiple Days in Ca<sup>2+</sup> Imaging Data****Graphical Abstract****Authors**

Liron Sheintuch, Alon Rubin,  
Noa Brande-Eilat, Nitzan Geva,  
Noa Sadeh, Or Pinchasof, Yaniv Ziv

**Correspondence**

yaniv.ziv@weizmann.ac.il

**In Brief**

Sheintuch et al. present a probabilistic method for tracking the same neurons across multiple days (cell registration) in large-scale Ca<sup>2+</sup> imaging data recorded from behaving mice. The probabilities for pairs of neighboring cells from different sessions to be the same neuron are estimated and utilized to perform reliable cell registration.

**Highlights**

- A method for tracking neurons across days (cell registration) in Ca<sup>2+</sup> imaging data
- The method is probabilistic and quantitatively evaluates registration accuracy
- The method is applicable to various imaging techniques and cell detection algorithms
- Registration accuracy remains high with an increased number of registered sessions



# Tracking the Same Neurons across Multiple Days in $\text{Ca}^{2+}$ Imaging Data

Liron Sheintuch,<sup>1</sup> Alon Rubin,<sup>1</sup> Noa Brande-Eilat,<sup>1</sup> Nitzan Geva,<sup>1</sup> Noa Sadeh,<sup>1</sup> Or Pinchasof,<sup>1</sup> and Yaniv Ziv<sup>1,2,\*</sup>

<sup>1</sup>Department of Neurobiology, Weizmann Institute of Science, Rehovot 76100, Israel

<sup>2</sup>Lead Contact

\*Correspondence: yaniv.ziv@weizmann.ac.il

<https://doi.org/10.1016/j.celrep.2017.10.013>

## SUMMARY

$\text{Ca}^{2+}$  imaging techniques permit time-lapse recordings of neuronal activity from large populations over weeks. However, without identifying the same neurons across imaging sessions (cell registration), longitudinal analysis of the neural code is restricted to population-level statistics. Accurate cell registration becomes challenging with increased numbers of cells, sessions, and inter-session intervals. Current cell registration practices, whether manual or automatic, do not quantitatively evaluate registration accuracy, possibly leading to data misinterpretation. We developed a probabilistic method that automatically registers cells across multiple sessions and estimates the registration confidence for each registered cell. Using large-scale  $\text{Ca}^{2+}$  imaging data recorded over weeks from the hippocampus and cortex of freely behaving mice, we show that our method performs more accurate registration than previously used routines, yielding estimated error rates <5%, and that the registration is scalable for many sessions. Thus, our method allows reliable longitudinal analysis of the same neurons over long time periods.

## INTRODUCTION

Recent advances in optical imaging techniques and genetically encoded  $\text{Ca}^{2+}$  indicators allow researchers to chronically record the activity of hundreds to thousands of neurons simultaneously in behaving animals (Svoboda et al., 1997; Helmchen et al., 2001; Sawinski et al., 2009; Ghosh et al., 2011; Grienberger and Konnerth, 2012; Chen et al., 2013). This is typically done using two-photon imaging in head-fixed rodents (Domek et al., 2007, 2010; Kitamura et al., 2015; Burgess et al., 2016) or one-photon imaging with miniature microscopes in freely behaving rodents (Ziv et al., 2013; Berdyyeva et al., 2014; Jennings et al., 2015; Pinto and Dan, 2015; Sun et al., 2015; Cai et al., 2016). These techniques facilitate within-subject analyses that quantify changes in neuronal activity under different experimental conditions and over extended periods of time (reviewed in Ziv and Ghosh, 2015). However, without identifying the same individual neurons across imaging sessions, the analysis of

time-lapse imaging data becomes limited to population-level statistics, thus losing a critical advantage offered by optical imaging relative to dense electrophysiological recordings. Following the activity of the same neurons over time can uncover the changes in the coding properties of individual cells and in the joint activity patterns that underlie the population-level statistics (Huber et al., 2012; Ziv et al., 2013; Lütcke et al., 2013; Jennings et al., 2015; Rubin et al., 2015; Burgess et al., 2016; Cai et al., 2016; Liberti et al., 2016; Rose et al., 2016; Driscoll et al., 2017; Grewe et al., 2017). Such accounts of long-term dynamics are crucially absent in many fields of neuroscience, including the study of learning and long-term memory, where understanding how information is represented, stored, and changes with time is key.

To longitudinally follow the activity of individual neurons, the same cells need to be reliably identified (registered) across all time points in the experiment. Registration of the same neurons becomes challenging as the number of detected cells in a session, the number of sessions, and the intervals between them increase. Cell registration is further complicated in data from one-photon  $\text{Ca}^{2+}$  imaging, in comparison to two-photon imaging, due to light scattering and lack of optical sectioning (Wilt et al., 2009; Yang and Yuste, 2017), both of which increase the crosstalk between the signals of neighboring neurons in the two-dimensional field of view (FOV). Furthermore, since one-photon microscopy mostly reveals transient and localized changes in fluorescence that exceed the background noise, a cell must be active to be detected. Thus, the set of detected cells can differ between sessions (Ziv et al., 2013; Resendez et al., 2016; Grewe et al., 2017; Xia et al., 2017), and a one-to-one mapping of neural identity across sessions is typically not attainable. Taken together, these factors introduce uncertainty and, consequently, potential errors to the cell registration procedure.

While numerous methods for detecting cells and extracting their activity from  $\text{Ca}^{2+}$  imaging data have been developed (Reidl et al., 2007; Vogelstein et al., 2009, 2010; Mukamel et al., 2009; Grewe et al., 2010; Smith and Häusser, 2010; Oñativia et al., 2013; Pachitariu et al., 2013; Maruyama et al., 2014; Pnevmatikakis et al., 2016; Theis et al., 2016), relatively little effort has been devoted to the issue of cell registration across sessions. Previous studies that registered cells across sessions (Ziv et al., 2013; Jennings et al., 2015; Rubin et al., 2015; Cai et al., 2016; Liberti et al., 2016; Kitamura et al., 2017), regardless of whether using manual or automatic routines, did not provide a quantitative evaluation of registration accuracy in terms of false-positive errors (different cells falsely registered as the

same cells) and false-negative errors (the same cells falsely registered as different cells). The lack of such quantitative evaluation could be detrimental in cases where the acquired data are inadequate for longitudinal analysis and may lead to misinterpretation of the data (Harris et al., 2016). For example, either falsely identifying two different cells as the same or falsely identifying the same cell as two different cells can lead to false conclusions about the dynamics of the neuronal activity. Moreover, previous work has used fixed registration decision parameters (e.g., distance threshold), which were not optimized to the specific data.

To address these problems, we adopted a probabilistic approach to devise a method for automated cell registration across sessions. By modeling the distribution of similarities between neighboring cells across sessions, our method estimates the confidence of registration associated with each cell in the data. Moreover, it estimates the overall rates of false-positive and false-negative errors for different registration decision parameters. This approach enables cell registration that is adaptive and optimized to different datasets. Applying our method to data recorded from the hippocampus and cortex of behaving mice, we show that the same cells can be tracked over multiple weeks with estimated false-positive and false-negative rates <5%, yielding more accurate registration than the routines utilized in previous studies. Moreover, we find that registration accuracy remains high with increased numbers of sessions, demonstrating the method's suitability for longitudinal studies. We provide an open source MATLAB code for cell registration that implements the approach presented in this paper (see *Experimental Procedures*).

## RESULTS

### Cells Maintain Their Spatial Footprints across Multiple Days of Imaging

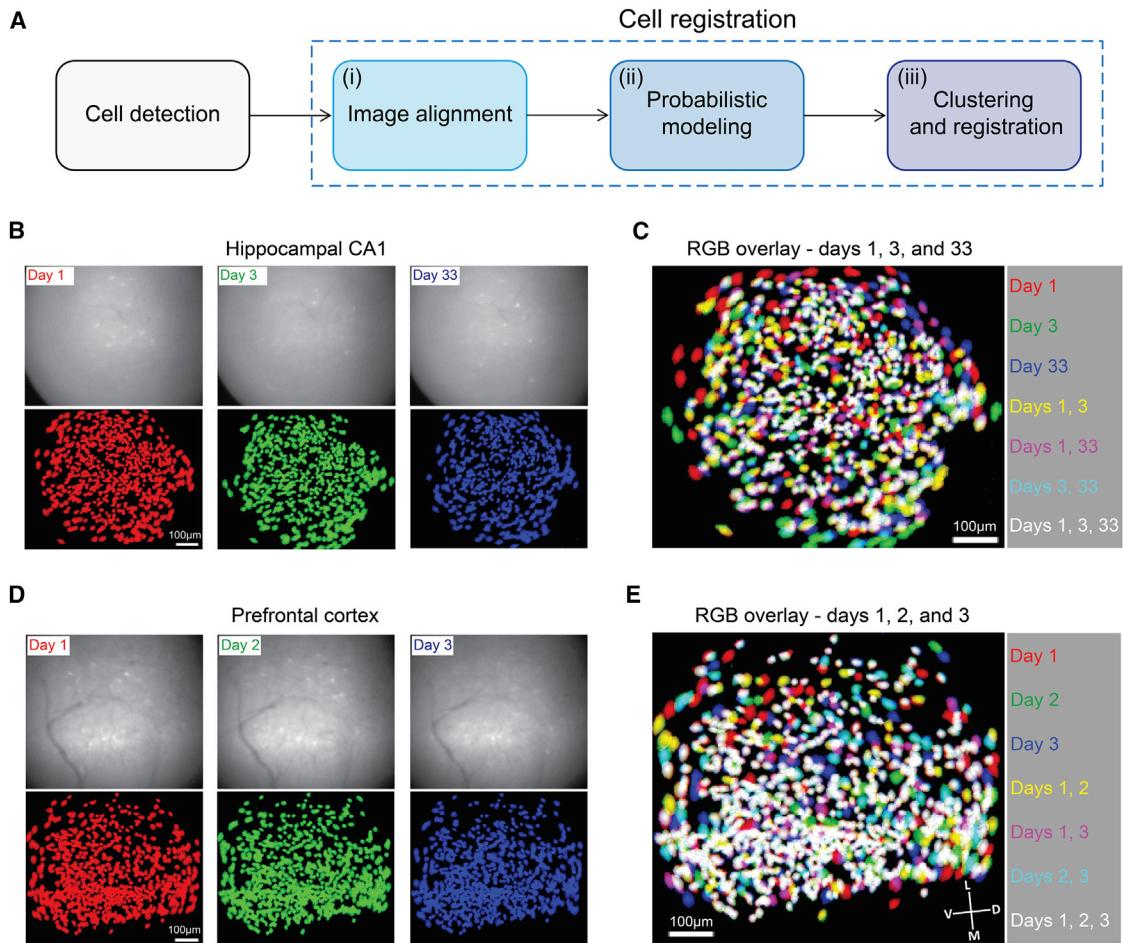
We combined miniaturized head-mounted microscopes (Ghosh et al., 2011) and microendoscopic probes (Barreto et al., 2011) to chronically image  $\text{Ca}^{2+}$  dynamics in GCaMP6-expressing neurons in the hippocampal CA1 and in the pre-frontal cortex of freely behaving mice. For each mouse, we imaged  $\text{Ca}^{2+}$  dynamics in multiple recording sessions over weeks (12 mice, 6–16 sessions per mouse, 269–712 cells per session). Prior to cell registration, cells that were active in each imaging session were detected, and their spatial footprints (i.e., weighted regions of interest consisting of each pixel's contribution to the cell's fluorescence) and  $\text{Ca}^{2+}$  traces were extracted (Figure S1) using an established routine based on principal-component analysis and independent-component analysis (PCA-ICA; Mukamel et al., 2009). To register cells across the different sessions, we constructed a cell registration method that consists of three main steps (Figure 1A): (1) aligning between the FOVs imaged in different sessions; (2) modeling the distribution of similarities between pairs of neighboring cells from different sessions to obtain an estimation for their probability to be the same cell; and (3) registering cells across multiple sessions via a clustering procedure that uses the obtained probabilities of neighboring cell-pairs to be the same cell.

To correct for translation and rotation differences between sessions, we aligned the FOV of each session with the FOV of a reference session, yielding the locations of spatial footprints from different sessions in a single coordinate system (Figures 1B–1E and S2). The cells generally maintained their spatial footprints over long time periods, as indicated by the overlap of spatial footprints across sessions.

### Spatial Footprint Similarities across Sessions Exhibit a Bimodal Distribution

We considered all pairs of cells that were detected in close proximity in the FOV across different sessions (*neighboring cell-pairs*) to be same cell candidates. For each neighboring cell-pair, we measured: (1) the distance between the centers of mass of their spatial footprints (*centroid distance*) and (2) the Pearson correlation between their spatial footprints (*spatial correlation*). While some of the neighboring cell-pairs had either large centroid distances and small spatial correlations (suggesting that they are different cells) or small centroid distances and large spatial correlations (suggesting that they are the same cells), for other cell-pairs, intermediate values were observed, indicating that their registration was uncertain (Figure 2A). Therefore, we sought to register cells across sessions while obtaining a quantitative estimation of registration confidence. We limited the analysis to cell-pairs with centroid distances <12  $\mu\text{m}$ , assuming that cell-pairs with larger centroid distances are different cells. We then calculated the distributions of centroid distances and spatial correlations between nearest neighbors and between other (not nearest) neighbors across sessions (Figures 2B, 2C, and S3). Based on data from 12 mice,  $87\% \pm 3\%$  of the nearest neighbors had a centroid distance <7  $\mu\text{m}$ , and  $89\% \pm 4\%$  had a spatial correlation >0.6, while only  $5\% \pm 1\%$  of the other neighbors had a centroid distance <7  $\mu\text{m}$ , and  $6\% \pm 2\%$  had a spatial correlation >0.6. The differences between the distributions for nearest neighbors and other neighbors support the notion that nearest neighbors are mostly the same cells, while other neighbors are, for the most part, different cells. However, registering all pairs of nearest neighbors as the same cells would result in false-positives when a cell is active in only one of the two sessions, as indicated by the heavy tail in the distributions for nearest neighbors. Furthermore, since the distributions for nearest neighbors and other neighbors partially overlap, any registration threshold, i.e., a value that serves as a cutoff for deciding whether two cells are the same, would result in false-positive errors, false-negative errors, or both.

To estimate the attainable accuracy of cell registration, we calculated the distributions of centroid distances and spatial correlations between all (nearest and other) neighboring cell-pairs from different sessions (Figures 2D and 2E). Both distributions were bimodal, reflecting the fact that they comprise two different subpopulations of cell-pairs corresponding to same cells and different cells. Therefore, we modeled the data (centroid distances model and spatial correlations model) as a weighted sum of the distributions of two subpopulations. For each mouse, we jointly estimated the parameters of the distributions of the two subpopulations and their weights to best fit the data (Figures 2D, 2E, and S4A–S4F, showing additional examples for mice



**Figure 1. Cells Maintain Their Locations and Shapes over Weeks**

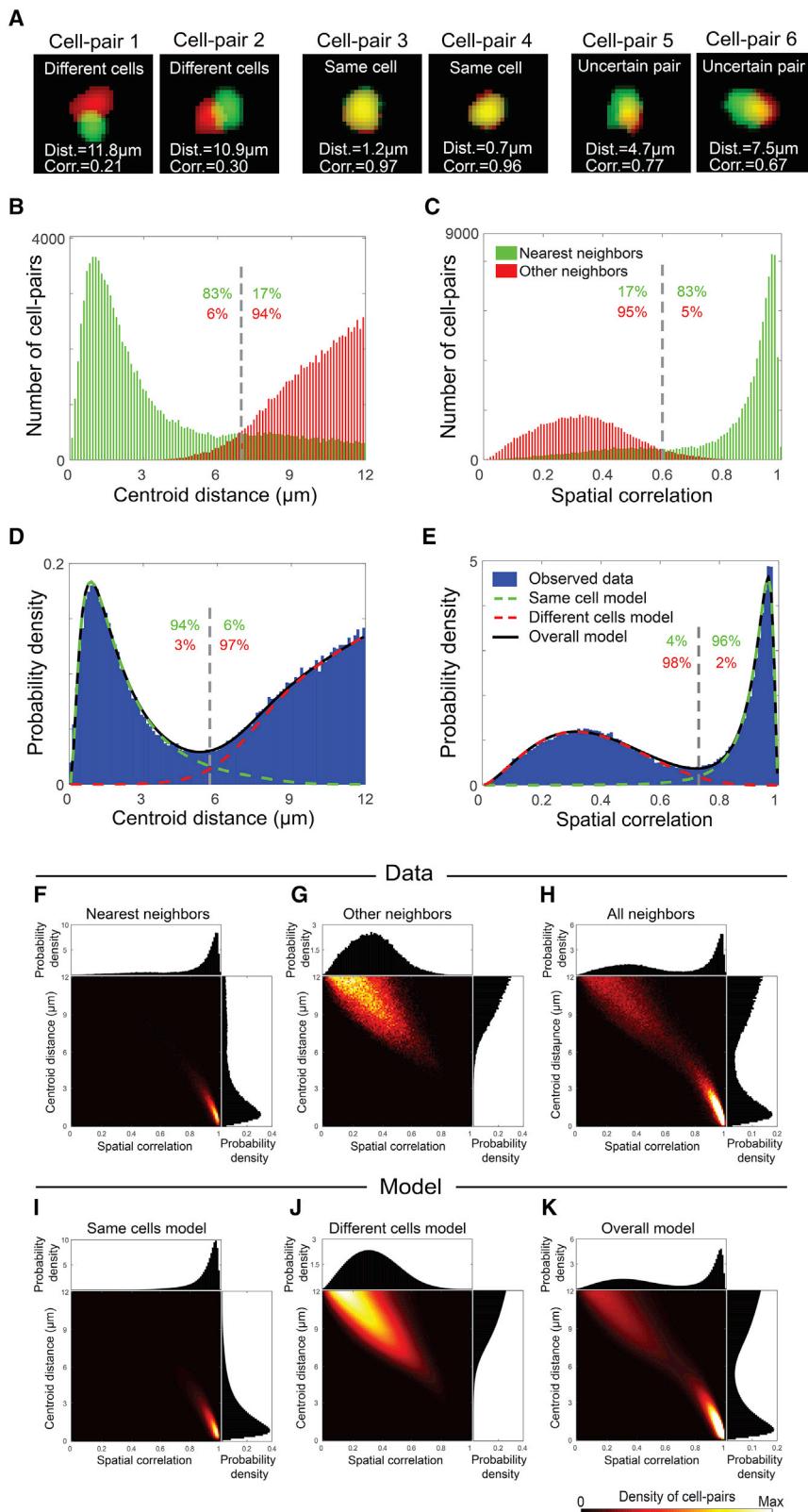
(A–E) In (A), the main steps in the cell registration procedure are indicated. (B and D) Top: representative single frames from raw fluorescence data of imaging sessions recorded on three different days. Bottom: projection of all spatial footprints for the same three sessions, indicated in red, green, and blue. (B) Hippocampal CA1. (D) Prefrontal cortex. (C and E) Overlays of the aligned spatial footprint maps shown for (C) hippocampal CA1, as shown in (B), and for (E) prefrontal cortex, as shown in (D). D, dorsal; L, lateral; M, medial; V, ventral. Data were recorded in the hippocampal CA1 of a Thy1-GCaMP6f transgenic mouse (B and C) and in the prefrontal cortex of a CaMKII-GCaMP6s transgenic mouse (D and E) while freely exploring the same environments. See also Figures S1 and S2.

imaged in the prefrontal cortex or hippocampus). The obtained estimated distributions enable an informed choice of the registration threshold that is optimized to the data of each mouse (Figures 2D and 2E, gray dashed lines at the intersection of the two distributions; Figures S4G and S4H). Similarly, we used the joint distribution of centroid distances and spatial correlations across sessions to model the data (joint model; Figures 2F–2K, S5, and S6).

#### A Large Fraction of Cell-Pairs Can Be Reliably Registered across Sessions

The resulting models allowed us to use a Bayesian framework to calculate the probability for two cells from two different sessions to be the same cell ( $P_{\text{same}}$ ), given their spatial correlation and centroid distance (Figures 3A and 3B). By applying  $P_{\text{same}}$  (calculated for each mouse individually) to the distribution of cell-pairs in our data (802,862 neighboring cell-pairs pooled from 12 mice),

we obtained the cumulative fraction of cell-pairs as a function of  $P_{\text{same}}$  (Figure 3C). We defined uncertain registrations of cell-pairs as those with a probability  $\leq 0.95$  to be the same cells and a probability  $\leq 0.95$  to be different cells ( $0.05 \leq P_{\text{same}} \leq 0.95$ ) and found that the spatial correlations and joint models result in less uncertainty than the centroid distances model (inset of Figure 3C;  $19.7\% \pm 3.3\%$ ,  $16.6\% \pm 5.5\%$ , and  $15.1\% \pm 3.4\%$  of the neighboring cell-pairs are in the uncertain range for the centroid distances, spatial correlations, and joint models, respectively; one-way repeated-measures ANOVA,  $F(1, 16) = 11.6$ ,  $p = 0.002$ ). Additionally, we used the estimated probability distributions of same cells and different cells to calculate a receiver operating characteristic (ROC) curve, providing the estimated false-positive and false-negative rates for various registration thresholds (Figure 3D). We calculated the Gini coefficient  $G_1$  (a measure for the area under the curve) and found that the joint model performs slightly better than the centroid



**Figure 2. Distributions of Spatial Footprint Similarities Modeled as a Weighted Sum of Two Subpopulations**

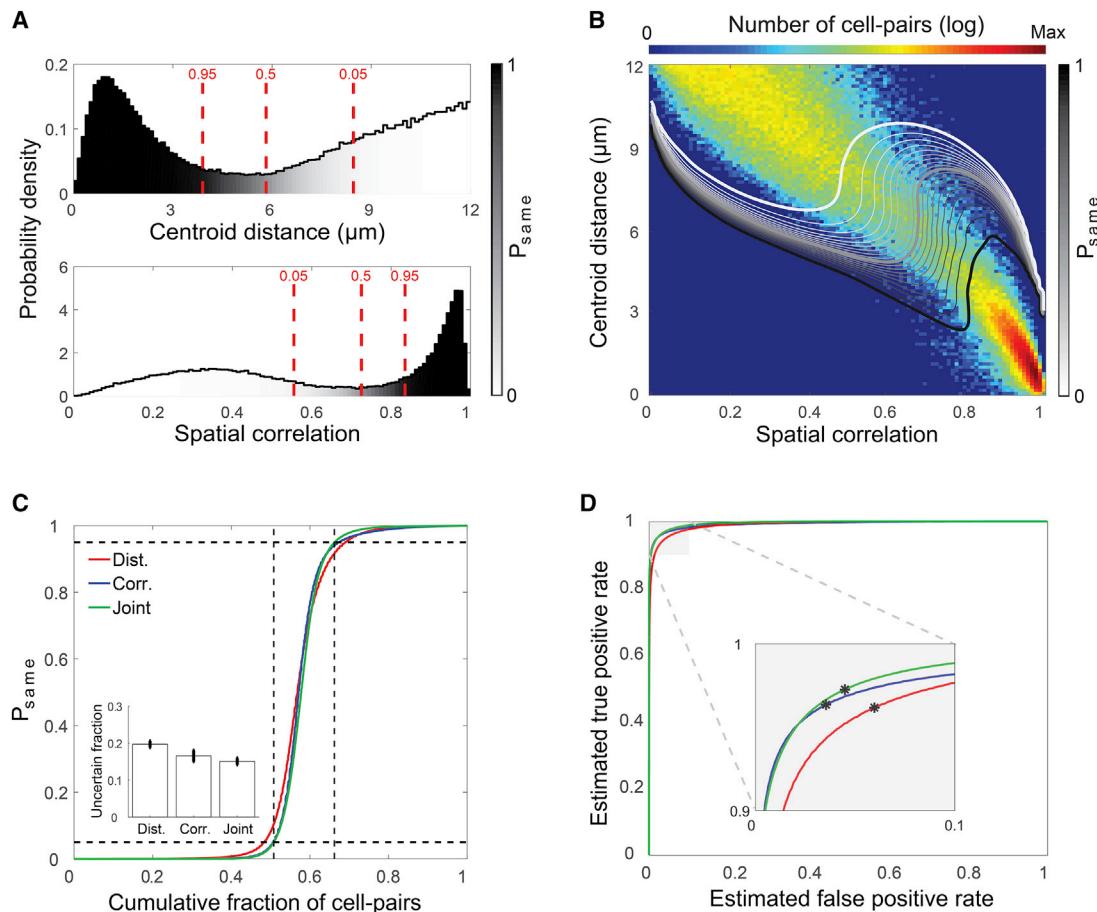
(A) Six examples of candidates to be the same cell, with their measured centroid distances (Dist.) and spatial correlations (Corr.). The spatial footprints are shown in red (session 1) and green (session 2). (B) and (C) Distribution of centroid distances (B) and spatial correlations (C) between pairs of nearest neighbors (green) and other neighbors (red) across sessions. Gray dashed lines show the intersection between the two distributions. The fraction of nearest neighbors (green) or other neighbors (red) above and below these intersections are indicated.

(D) and (E) Distributions of centroid distances (D) and spatial correlations (E) between all neighboring cell-pairs (blue bars) and the modeled distributions of same cells (dashed green curves), different cells (dashed red curves), and their weighted sum (solid black curves). Gray dashed lines show the intersection between the two models. Estimated fractions of same cells (green) or different cells (red) above and below these intersections are indicated.

(F–H) Joint and marginal distributions of centroid distances and spatial correlations between pairs of nearest neighbors (F), other neighbors (G), and all neighbors (H) across sessions.

(I–K) Modeled joint and marginal distributions of centroid distances and spatial correlations for same cells (I), different cells (J), and their weighted sum (K). In (H) and (K), the color scale was set to reach 0.25 of the maximal value to enable visualization of both subpopulations. Data and models in all panels are for 16 sessions recorded on 8 different days in the hippocampal CA1 of a mouse while freely exploring the same environments.

See also Figures S3, S4, S5, and S6.



**Figure 3. Modeling the Data Supports Reliable Cell Registration across Pairs of Sessions**

(A and B) Shown in (A): distributions of centroid distances (top) and spatial correlations (bottom) between neighboring cell-pairs. The grayscale color code indicates the probability for two cells from two different sessions to be the same cell ( $P_{\text{same}}$ ). Red lines show  $P_{\text{same}} = 0.05$ ,  $P_{\text{same}} = 0.5$ , and  $P_{\text{same}} = 0.95$ . (B) Contours of  $P_{\text{same}}$  overlaid on the joint distribution of centroid distances and spatial correlations for neighboring cell-pairs. Numbers of cell-pairs are displayed in a logarithmic scale. The  $P_{\text{same}} = 0.05$ ,  $P_{\text{same}} = 0.5$ , and  $P_{\text{same}} = 0.95$  contours are highlighted by thicker white, gray, and black curves, respectively. Data in (A) and (B) are the same as in Figure 2.

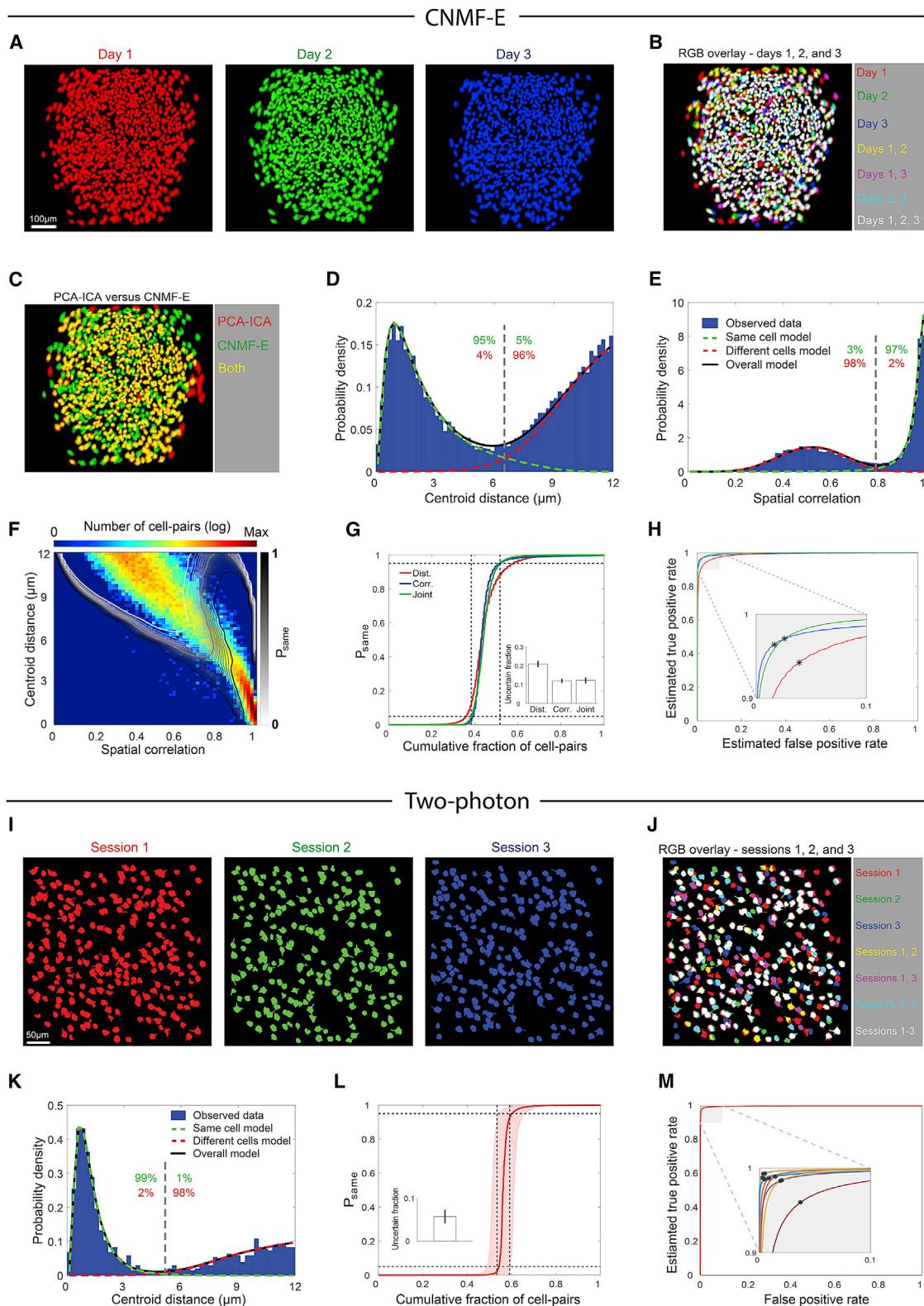
(C and D) Shown in (C): the cumulative fraction of cell-pairs as a function of the estimated  $P_{\text{same}}$  for the centroid distances (red), spatial correlations (blue), and joint (green) models. Black dashed horizontal lines represent the  $P_{\text{same}} = 0.05$  and  $P_{\text{same}} = 0.95$  levels, while the vertical lines show the joint model's intersection with those probabilities, indicating the fraction of uncertain registrations. Inset: the fraction of cell-pairs in the uncertain registration range (mean  $\pm$  SEM). (D) Estimated ROC curves for the centroid distances (red), spatial correlations (blue), and joint (green) models. Inset: zoom-in on the near-optimal part of the ROC. Black asterisks represent the  $P_{\text{same}} = 0.5$  threshold. Data in (C) and (D) were pooled from 12 mice.

distances and spatial correlations models ( $G_1 = 0.983 \pm 0.005$ ,  $G_1 = 0.986 \pm 0.008$ , and  $G_1 = 0.991 \pm 0.004$  for the centroid distances, spatial correlations, and joint models, respectively; one-way repeated-measures ANOVA,  $F(1, 16) = 13.8$ ,  $p = 0.001$ ). For a threshold of  $P_{\text{same}} = 0.5$  (see black asterisks in the inset of Figure 3D), the false-negative rates were lower for the joint model ( $2.7\% \pm 0.9\%$ ) compared to the spatial correlations and the centroid distances models ( $3.7\% \pm 1.1\%$  and  $3.8\% \pm 1.1\%$ , respectively; one-way repeated-measures ANOVA,  $F(1, 14) = 10.9$ ,  $p = 0.003$ ), while the false-positive rates were lower for the spatial correlations model ( $3.6\% \pm 1.6\%$ ) compared to the centroid distances and joint models ( $6.0\% \pm 2.3\%$  and  $4.5\% \pm 1.9\%$ , respectively; one-way repeated-measures ANOVA,  $F(2, 19) = 28.4$ ,  $p < 0.001$ ). Overall, these results demonstrate that accurate registration across

sessions can be obtained in  $>95\%$  of neighboring cell-pairs based on any of the three models.

### Probabilistic Approach for Cell Registration Is Applicable to Different Types of Data

It has recently been shown that constrained nonnegative matrix factorization (CNMF) is well suited for extracting cellular spatial footprints and temporal dynamics from  $\text{Ca}^{2+}$  imaging data with significant spatial overlap between neighboring neural components (Pnevmatikakis et al., 2016). Therefore, in addition to the commonly used PCA-ICA method, we also used CNMF-E (Zhou et al., 2016), an extension of the CNMF method for one-photon microendoscopic data (Figures 4A and 4B). Notably, using CNMF-E, we detected within the same FOVs  $60.2\% \pm 36.5\%$  more cells than with PCA-ICA (five mice, five sessions

**Figure 4. The Cell Registration Method Is Applicable to Different Types of Imaging Data**

(A–H) Registration is applicable to cells detected using CNMF-E.

(A) Projection of all spatial footprints for three imaging sessions recorded on three different days, indicated in red, green, and blue.

(legend continued on next page)

per mouse, 348–1,498 cells per session) (Figure 4C). Applying our registration procedure to the detected cells, we obtained a probabilistic model of the data without the need for any adaptations to our method (Figures 4D–4F). Although we detected a higher density of cells with CNMF-E, we could, nonetheless, register cells (198,230 neighboring cell-pairs) with certainty levels, false-negative rates, and false-positive rates comparable to those obtained with PCA-ICA (Figures 4G and 4H).

In addition, we tested our registration method on data from two-photon imaging in head-fixed behaving mice (10 mice, three sessions per mouse, 196–375 cells per session, obtained from the Allen Brain Observatory, 2016). Image alignment of the different sessions revealed that different subsets of cells were detected in each of the imaging sessions (Figures 4I and 4J), demonstrating that there was no one-to-one mapping of cell identity across sessions, consistent with previous reports (Driscoll et al., 2017; Wagner et al., 2017). As expected from the optical sectioning and higher spatial resolution of two-photon microscopy, cell registration (19,970 neighboring cell-pairs) yielded lower registration uncertainty (Figures 4K and 4L;  $5.8\% \pm 4.8\%$  of the neighboring cell-pairs are in the uncertain range), false-negative rates ( $1.4\% \pm 0.8\%$ ), and false-positive rates (Figure 4M;  $1.3\% \pm 1.3\%$ ) than those obtained with one-photon microscopy. Overall, our registration method was found to be applicable to both one-photon and two-photon imaging techniques, to various cell detection algorithms, and to imaging data from different brain regions (Figures S4 and S5).

### Cell Registration Is Scalable for a Large Number of Sessions

Consistent with the observed bimodal distributions of cell-pair similarity (Figure 2), the centroid locations of cells across multiple sessions seem to be spatially clustered (Figure 5A). Consequently, high  $P_{\text{same}}$  values are expected between cell-pairs within a cluster, whereas low  $P_{\text{same}}$  values are expected between cell-pairs from different clusters. Therefore, to register cells from

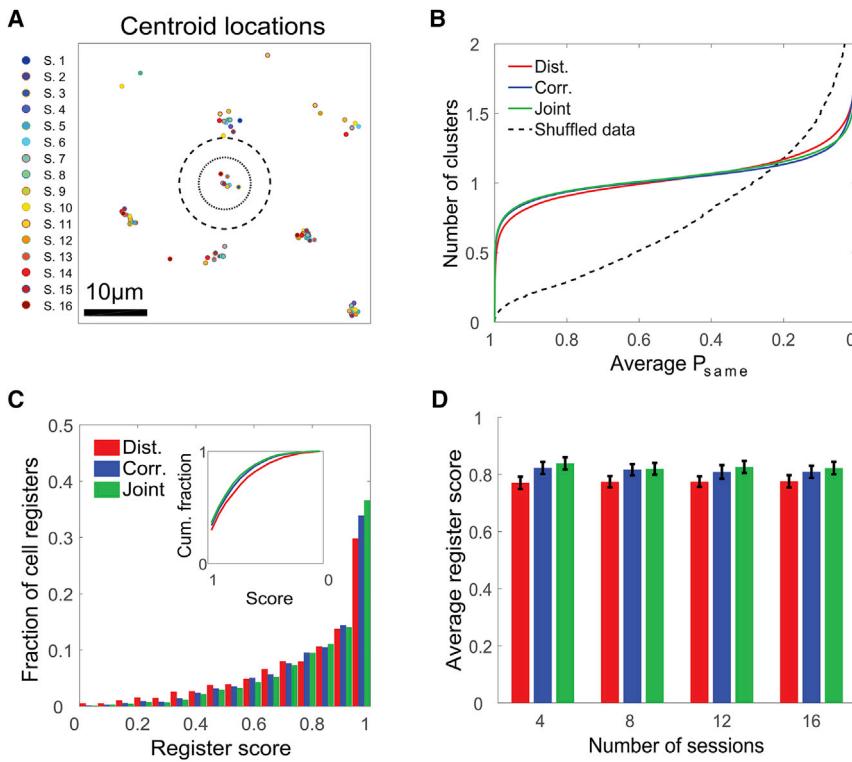
multiple sessions as single entities, we applied a clustering procedure (Bansal et al., 2004) based on the estimated  $P_{\text{same}}$ . As expected, clustering of the cells from all the sessions revealed that each cell tends to have a high  $P_{\text{same}}$  with one and only one cluster (Figure 5B).

The clustering procedure yields a list of entities (cell registers), each representing a single cell over multiple sessions. To quantify the registration quality for each cell register, we defined a stringent register score index (range = 0–1), which takes into account the registration certainty of all cell-pairs within the cluster (Figures 5C and S7). Register scores (10,102 cell registers pooled from 12 mice) were higher for the joint model relative to the centroid distances model (average register score of  $0.78 \pm 0.02$ ,  $0.81 \pm 0.02$ , and  $0.82 \pm 0.02$ , for the centroid distances, spatial correlations, and joint models, respectively; repeated-measures ANOVA,  $F(1, 5) = 10.8$ ,  $p = 0.024$ ). To test whether cell registration accuracy is maintained over a large number of sessions over weeks, we used data recorded across 16 sessions spanning 15 days (Rubin et al., 2015) and repeatedly applied our registration method to various numbers of sessions. We did not find the average register score (4,242 cell registers pooled from five mice) to decrease with the number of registered sessions (repeated-measures ANOVA,  $F(2, 9) = 2.3$ ,  $p = 0.129$ ; Figure 5D), demonstrating that our registration method is scalable for a large number of sessions, rendering it suitable for longitudinal studies.

### Validation of the Cell Registration Method

We applied several independent techniques to validate cell registration across multiple sessions, both at the cell-pair level and at the cell-register level. First, we quantified *exclusivity* and *transitivity*, two logical prerequisites for the model to be internally consistent with the data. Exclusivity requires that, if a given cell from session 1 is paired with a cell from session 2, then no other cell from session 2 be paired with it. Transitivity requires that, if a cell from session 1 and a cell from session 2 are paired, and the cell from session 2 is paired with a cell from session 3,

- (B) Overlay of the aligned spatial footprint maps shown in (A).
- (C) Overlay of spatial footprint maps detected using PCA-ICA (red) and CNMF-E (green) for the same session.
- (D and E) Distributions of centroid distances (D) and spatial correlations (E) between all neighboring cell-pairs (blue bars) and the modeled distributions of same cells (dashed green curves), different cells (dashed red curves), and their weighted sum (solid black curves). Gray dashed lines show the intersection between the two distributions. Estimated fractions of same cells (green) or different cells (red) above and below these intersections are indicated.
- (F) Contours of  $P_{\text{same}}$  overlaid on the joint distribution of centroid distances and spatial correlations for neighboring cell-pairs. Numbers of cell-pairs are displayed in a logarithmic scale. The  $P_{\text{same}} = 0.05$ ,  $P_{\text{same}} = 0.5$ , and  $P_{\text{same}} = 0.95$  contours are highlighted by thicker white, gray, and black curves, respectively.
- (G) The cumulative fraction of cell-pairs as a function of the estimated  $P_{\text{same}}$  for the centroid distances (red), spatial correlations (blue), and joint (green) models. Black dashed horizontal lines represent the  $P_{\text{same}} = 0.05$  and  $P_{\text{same}} = 0.95$  levels, while the vertical lines show the joint model's intersection with those probabilities. Inset: the fraction of cell-pairs in the uncertain registration range (mean  $\pm$  SEM).
- (H) Estimated ROC curves for the centroid distances (red), spatial correlations (blue), and joint (green) models. Inset: zoom-in on the near-optimal part of the ROC. Black asterisks represent the  $P_{\text{same}} = 0.5$  threshold. Data in (G) and (H) were pooled from five mice.
- (I–M) Registration is applicable to two-photon imaging data.
- (I) Projection of all spatial footprints for three imaging sessions recorded on three different days, indicated in red, green, and blue.
- (J) Overlay of the aligned spatial footprint maps shown in (I).
- (K) Distribution of centroid distances for neighboring cell-pairs (blue bars) and the modeled distribution of same cells (dashed green curve), different cells (dashed red curve), and their weighted sum (solid black curve).
- (L) Cumulative fraction of cell-pairs as a function of the estimated  $P_{\text{same}}$  for the centroid distances model (average  $\pm$  SD represented by the red curve and shaded red area). Black dashed horizontal lines represent the  $P_{\text{same}} = 0.05$  and  $P_{\text{same}} = 0.95$  levels, while the vertical lines show the model's intersection with those probabilities. Inset: the fraction of cell-pairs in the uncertain registration range (mean  $\pm$  SEM).
- (M) Average estimated ROC curve. Inset: zoom-in on the near-optimal part of the ROC obtained for each mouse individually. Black asterisks represent the  $P_{\text{same}} = 0.5$  threshold. Two-photon data in (I)–(M) were obtained from the Allen Brain Observatory, 2016, based on experiments consisting of three sessions recorded on three different days in the visual cortex of head-fixed behaving mice. Data in (L) and (M) were pooled from 10 mice.



**Figure 5. Accurate Cell Registration Is Scalable for a Large Number of Sessions**

(A) Centroid locations of cells from 16 different sessions taken from a small part of the FOV. The inner circle represents a radius of 4  $\mu\text{m}$ , and the outer (dashed) circle represents a radius of 7  $\mu\text{m}$ .

(B) The number of candidate clusters per cell remains  $\sim 1$  over a wide range of average  $P_{\text{same}}$  values for the centroid distances (red), spatial correlations (blue), and joint (green) models but not for the shuffled data (black dashed curve).

(C) Distribution of register scores for the centroid distances (red), spatial correlations (blue), and joint (green) models. Inset: cumulative fraction of cell registers as a function of the register score reversed from 1 to 0.

(D) Register score as a function of the number of registered sessions for the centroid distances (red), spatial correlations (blue), and joint (green) models (mean  $\pm$  SEM).

Data in (B) and (C) were pooled from 12 mice. Data in (D) were pooled from five mice.

See also Figure S7.

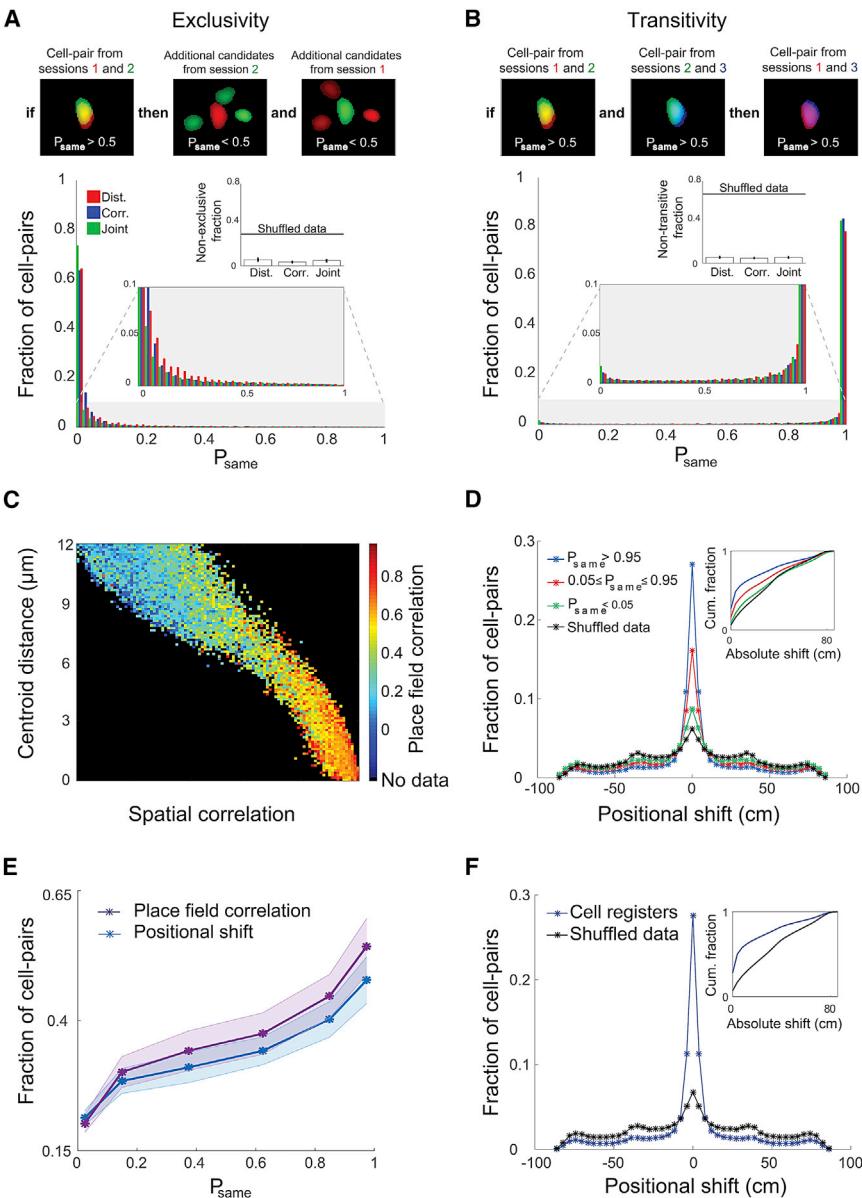
then the cells from sessions 1 and 3 also be paired. To measure exclusivity, we focused on cells that have a pairing candidate with  $P_{\text{same}} > 0.5$  in another session and calculated for these cells the distribution of  $P_{\text{same}}$  with all additional pairing candidates (Figure 6A). As expected, the additional pairing candidates had a low  $P_{\text{same}}$  with  $93.9\% \pm 4.5\%$ ,  $96.1\% \pm 2.4\%$ , and  $94.8\% \pm 3.6\%$  of them having a  $P_{\text{same}} < 0.5$  for the centroid distances, spatial correlations, and joint models, respectively, compared to 69.1% in the shuffled data (inset of Figure 6A). The joint model resulted in higher exclusivity compared to the centroid distances model (one-way repeated-measures ANOVA,  $F(1, 13) = 5.7$ ,  $p = 0.03$ ). To measure transitivity, we calculated the distribution of  $P_{\text{same}}$  for pairs of cells from two different sessions where each had a  $P_{\text{same}} > 0.5$  with the same candidate from a third session (Figure 6B). Indeed, such pairs showed a high  $P_{\text{same}}$  with  $94.1\% \pm 1.7\%$ ,  $94.9\% \pm 1.2\%$ , and  $94.3\% \pm 1.5\%$  of them having a  $P_{\text{same}} > 0.5$  for the centroid distances, spatial correlations, and joint models, respectively, compared to 32.8% in the shuffled data (inset of Figure 6B). We found no differences in transitivity between the three different models.

For an independent validation of the probabilistic approach, we applied our registration method to data recorded from hippocampal CA1 cells in mice that explored linear tracks (Rubin et al., 2015), focusing our analysis on place cells. Place cells, which are abundant in the hippocampus, selectively fire when an animal moves through a particular location in space (O'Keefe and Burgess, 1971). Previous work has shown that place cells that are active on different days tend to maintain their coding specificity, i.e., display stable place fields (Thompson and Best, 1990; Kennerley et al., 2004; Ziv et al., 2013). Therefore, cell-pairs with a high

$P_{\text{same}}$  are presumed to display more similar place fields across days than expected by chance (see Movie S1 for examples of place cells).

Cell-pairs (158,488 pairs of place cells pooled from five mice) with high spatial correlations and low centroid distances demonstrated high place-field correlations (Figure 6C), consistent with the notion that they are the same cell. We divided cell-pairs into three subsets: (1) high probability of being the same cell ( $P_{\text{same}} > 0.95$ ); (2) uncertain registration range ( $0.05 \leq P_{\text{same}} \leq 0.95$ ); and (3) high probability of being different cells ( $P_{\text{same}} < 0.05$ ). We calculated for each subset the place-field positional shifts and place-field correlations across days (Figures 6D and 6E) and compared them to the distribution obtained for shuffled data. As expected, cell-pairs with a higher  $P_{\text{same}}$  had higher probabilities to display minor positional shifts and high place-field correlations across days (48.8%, 33.1%, and 21.4% of cell-pairs had a place-field positional shift  $\leq 6$  cm for  $P_{\text{same}} > 0.95$ ,  $0.05 \leq P_{\text{same}} \leq 0.95$ , and  $P_{\text{same}} < 0.05$ , respectively, compared to 15.9% in shuffled data). Additionally, to validate our clustering procedure, we measured the place-field stability of cell registers across multiple sessions (Figure 6F). Cell registers (3,531 place cell registers pooled from five mice across eight sessions) had considerably more stable place fields across days than shuffled data (50.0% of the pairs within cell registers had a place-field positional shift  $\leq 6$  cm, compared to 17.0% for shuffled data), consistent with previous studies (Ziv et al., 2013; Rubin et al., 2015). The comparable levels of place-field stability observed for cell registers and for cell-pairs with  $P_{\text{same}} > 0.95$  imply that cells from multiple imaging sessions that were registered together, indeed, have a high probability to be the same neuronal entity.

Finally, we sought to use simulated data to validate our cell registration method across multiple sessions and evaluate its performance relative to previously used routines. We simulated



**Figure 6. Validation Based on the Exclusivity and Transitivity Principles and on Place-Field Stability across Days**

(A and B) Internal consistency between the model and the data. Exclusivity (A) and transitivity (B) measures for the centroid distances (red), spatial correlations (blue), and joint (green) models. (A) Top: illustration of the exclusivity principle. Bottom: distribution of  $P_{\text{same}}$  with additional pairing candidates, computed for cell-pairs from different sessions with  $P_{\text{same}} > 0.5$ . Insets: bottom, zoom-in on the lower area of the distribution; top, fraction of non-exclusive cell-pairs, i.e.,  $P_{\text{same}} > 0.5$  (mean  $\pm$  SEM). (B) Top: illustration of the transitivity principle. Bottom: distribution of  $P_{\text{same}}$  for cell-pairs from different sessions, where each had  $P_{\text{same}} > 0.5$  with the same cell from a third session. Insets: bottom, zoom-in on the lower area of the distribution; top, fraction of non-transitive cell-pairs, i.e.,  $P_{\text{same}} < 0.5$  (mean  $\pm$  SEM). Data in (A) and (B) were pooled from an N of 12 mice.

(C–F) Validation based on place-field stability across days.

(C) Average place-field correlation between cell-pairs across days given their centroid distances and spatial correlations. The black area represents centroid distances and spatial correlations with no cell-pairs.

(D) Distribution of place-field positional shifts across days for cell-pairs with  $P_{\text{same}} > 0.95$  (blue),  $0.05 \leq P_{\text{same}} \leq 0.95$  (red), and  $P_{\text{same}} < 0.05$  (green) and for shuffled data (black). Inset: the cumulative fraction of cell-pairs as a function of the absolute value of the positional shift.

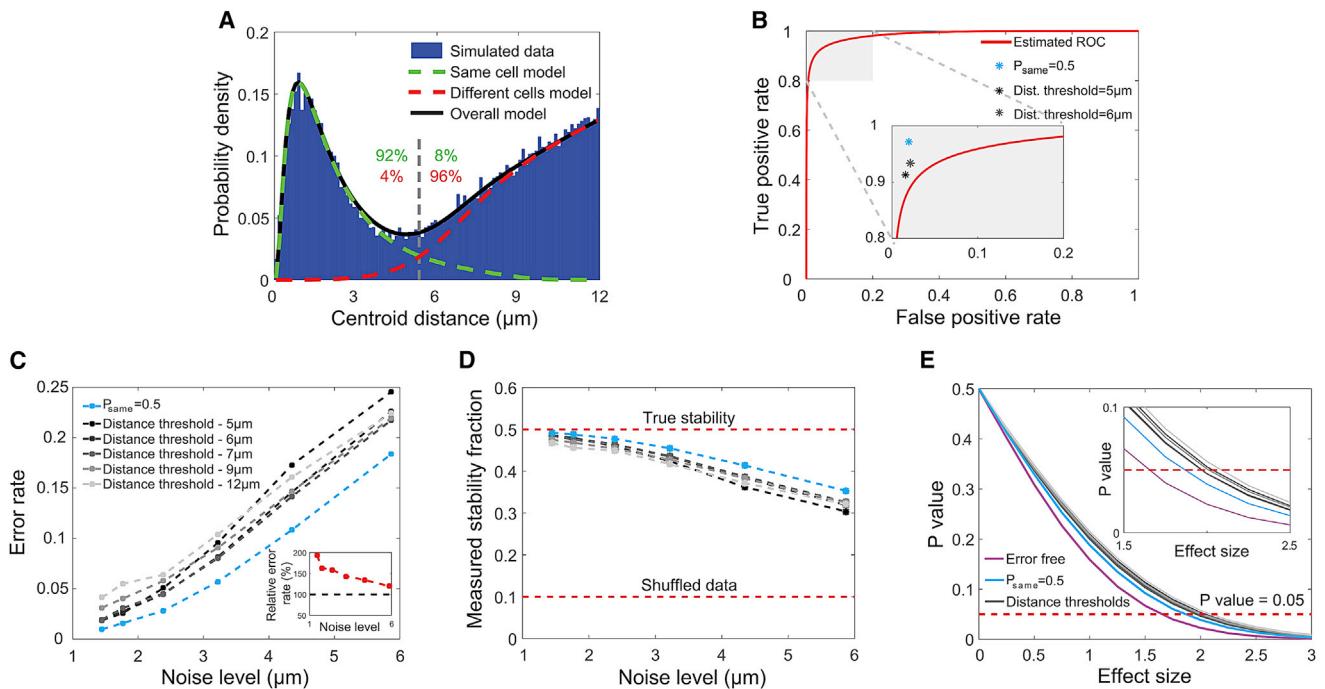
(E) Average fraction of neighboring cell-pairs with positional shifts  $\leq 6$  cm (blue) or place-field correlations  $> 0.5$  (purple) increase with  $P_{\text{same}}$  (linear regression:  $r^2 = 0.94$ ,  $p = 0.0015$ ; and  $r^2 = 0.93$ ,  $p = 0.0018$ , respectively; SD is indicated with a shaded area).

(F) Distribution of place-field positional shifts for cell-pairs taken from cell registers (blue) and shuffled data (black). Inset: the cumulative fraction of cell-pairs as a function of the absolute value of the positional shift. Data in (C)–(F) were pooled from an N of 5 mice.

See also Movie S1.

the centroid locations of cells across multiple sessions based on the distribution of centroid distances of same cells estimated from the data. Applying our registration procedure to the simulated data (Figure 7A), we obtained the estimated false-positive and false-negative rates (Figure 7B), as in Figure 3D. We then compared the actual false-positive and false-negative rates obtained with a registration threshold of  $P_{\text{same}} = 0.5$  to those obtained with previously used centroid distance thresholds of 5  $\mu\text{m}$  (Jennings et al., 2015; Cai et al., 2016) and 6  $\mu\text{m}$  (Ziv et al., 2013; Rubin et al., 2015). We found that our method yielded lower false-negative rates than registration routines used in previous studies (3.7% versus 8.6% and 6.8% for 5  $\mu\text{m}$  and 6  $\mu\text{m}$ , respectively) while maintaining comparable false-positive rates (1.9% versus 1.8% and 2.1% for 5  $\mu\text{m}$  and 6  $\mu\text{m}$ , respectively).

To study the performance of our method on imaging datasets of different quality levels, we simulated different noise levels and compared our registration method to registration using various centroid distance thresholds (Figure 7C). Our method yielded more accurate registrations across all noise levels and all registration thresholds. For the range of noise levels typically observed in one-photon imaging (1.5–3.5  $\mu\text{m}$ ; inset of Figure 7C), using the best of all tested thresholds for each noise level resulted in 43%–63% more registration errors than with our method. The decreased fraction of registration errors of our method yielded a more accurate estimation of neuronal coding stability (Figure 7D) and decreased the effect size required to expose coding stability that is significantly higher than chance levels (Figure 7E). While the effect size in this case (comparable

**Figure 7. Validation Based on Simulated Data**

(A) Distribution of centroid distances for neighboring cell-pairs (blue bars) and the modeled distribution of same cells (dashed green curve), different cells (dashed red curve), and their weighted sum (solid black curve). Estimated fractions of same cells (green) or different cells (red) above and below the intersection are indicated.

(B) Estimated ROC curve. Inset: zoom-in on the near-optimal part of the ROC. Asterisks represent the actual true-positive and false-positive rates obtained with a registration threshold of  $P_{\text{same}} = 0.5$  (blue), centroid distance of 5  $\mu\text{m}$  (black), and centroid distance of 6  $\mu\text{m}$  (dark gray).

(C) Error rate (weighted average of false-positive and false-negative rates) as a function of the noise level, for a registration threshold of  $P_{\text{same}} = 0.5$  (blue) and different centroid distance thresholds (shades of gray). Inset: the error rate obtained with the best centroid distance threshold (red dashed line) relative to the error rate obtained with a  $P_{\text{same}}$  value of 0.5 (black dashed line). Noise level is the mean centroid distance between the spatial footprints of the same cell in different sessions.

(D) Measured fractions of neurons with stable coding as a function of the noise level, for a registration threshold of  $P_{\text{same}} = 0.5$  (blue) and different centroid distance thresholds (shades of gray), compared to the true stability level and the stability fraction for shuffled data (red lines).

(E) The significance of the difference between the measured stability and the shuffled data as a function of the true population effect size, for registration with a threshold of  $P_{\text{same}} = 0.5$  (blue) and with different centroid distance thresholds (shades of gray), compared to an error-free measurement (magenta). Inset: zoom-in on the area with a  $P$  of 0.05. Effect size is the difference between the mean stability of the population and shuffled data divided by the stability SD.

For (A), (B), and (E), a noise level of 3.2  $\mu\text{m}$  was used.

to that observed experimentally for hippocampal place cells) is too large to be obscured by registration errors, other stability measures with smaller effect sizes may be more significantly impacted by registration accuracy. Notably, for different noise levels, different centroid distance thresholds yielded more accurate registration (Figures 7C and 7D). Since the optimal registration threshold is not known *a priori*, using a fixed centroid distance threshold decreases registration accuracy, further emphasizing the advantage of our method over previously used “one-size-fits-all” threshold-based registration.

## DISCUSSION

We investigated the ability to reliably track the same neurons across multiple weeks in  $\text{Ca}^{2+}$  imaging data and developed a method for automated cell registration across sessions. Because not all cells can be registered with certainty as same cells or as different cells, we used a probabilistic approach that

estimates the confidence of registration associated with each cell in the data. Using a Bayesian framework, we estimated the probability of neighboring cell-pairs from different sessions to be the same cell ( $P_{\text{same}}$ ), given their spatial correlation and centroid distance. All three tested models that are based on our probabilistic approach (centroid distances, spatial correlations, and joint) yielded high confidence levels of registration for a great majority of the cells (Figures 3C, 3D, and 5C). Importantly, we did not find registration accuracy to decrease as a function of the number of registered sessions (Figure 5D), suggesting that our method is scalable for a large number of sessions and demonstrating its utility for longitudinal studies.

We found that registration based on spatial correlations and on the joint models was slightly superior to that based on centroid distances. This result is consistent with the fact that spatial correlations account for the contribution of each pixel to the cell’s fluorescence, thus carrying additional information about cell identity that is absent in the previously used centroid

distance measure. Moreover, while optical aberrations (typical to gradient refractive index [GRIN] lens objectives used with miniature one-photon microscopes) may affect the measured centroid distance and result in larger distances in the periphery of the FOV, the spatial correlation is less sensitive to such artifacts.

Our method offers several measures to evaluate the suitability of the data for longitudinal studies. The measured distribution of spatial footprint similarities (either centroid distance or spatial correlation) across sessions constitutes one such measure: the greater the separation between the distributions for same cells and for different cells, the higher the accuracy of cell registration that can be obtained. Although this level of separation cannot be measured directly from the data, it is nevertheless reflected in the bimodality of the distribution for all neighboring cell-pairs across sessions (Figures 2D, 2E, and 2H). Our probabilistic approach is based on the quantification of this bimodality, providing the estimated false-positive and false-negative rates for different registration thresholds. These measures were lacking in previous works. This quantification also allows for adjusting the balance between conservativeness and permissiveness by controlling the tradeoff between false-positive and false-negative rates (Figure S7). For example, when studying the long-term stability of the coding properties of individual neurons, choosing a conservative registration threshold (e.g.,  $P_{\text{same}} = 0.95$ ) may be more appropriate (see the differences in place-field stability in Figure 6D). Additionally, the method quantifies the registration quality of each cell register in the data, which enables the inclusion, in the analysis, of only cell registers with high confidence levels. Finally, the method provides quantifications that can be used to evaluate the stability of the preparation (Figure S2) and detect abnormalities (e.g., changes in the focal plane over time). Taken together, these evaluations provide objective measures of data quality (Harris et al., 2016) that can help researchers determine whether their data are adequate for longitudinal analysis and to adaptively utilize the cell registration procedure to suit the specific needs of the study.

We validated our cell registration method using three independent analyses: (1) exclusivity and transitivity, which demonstrated internal consistency between the model and the data (Figures 6A and 6B); (2) coding stability of place cells that provided an unbiased validation of the method, since coding properties were not used for cell registration (Figures 6D–6F); and (3) registration of simulated data for which the ground truth is known, which demonstrated superior performance of our registration method over previously used routines that apply a predetermined centroid distance threshold (Figures 7B and 7C). This improved performance stems from two factors: (1) the probability model allows optimizing the registration threshold to the imaged dataset, and (2) the clustering procedure uses additional information from all neighboring cells across all sessions.

The method that we introduce here is particularly tailored for large-scale, one-photon imaging data, e.g., data collected using miniature fluorescence microscopes in freely behaving rodents. Indeed, for two-photon imaging data, there are alternative methods for cell registration that are based on image alignment

between a dynamic marker (e.g., GCaMP6) and a structural reference marker (e.g., a red fluorescent protein that is expressed in the nucleus; Rose et al., 2016), which enables one-to-one mapping between sessions. Such a method, however, cannot be used with one-photon imaging data due to lack of optical sectioning and light scattering, which preclude the resolution of fine structures (e.g., cell nucleus) from non-dynamic markers in dense labeling schemes.

We demonstrated the applicability of our registration method to data recorded from GCaMP6f/s-expressing neurons in the hippocampus and prefrontal cortex of both virus-injected and transgenic mice (Figures S4 and S5). We also found our registration procedure to be robust to various imaging techniques and cell detection algorithms: PCA-ICA and CNMF-E for one-photon data (Figures 1 and 4A–4H) and normalized periodical projection for two-photon data (Figures 4I–4M). While other types of data may require some adaptations to our method (e.g., the use of other similarity measures or functional forms to model the distribution of the data), the probabilistic approach we used is general and can accommodate such adaptations. The registration procedure is fully automated and does not require manual inspection of each cell, making it scalable for very large FOVs (Kim et al., 2016; Pachitariu et al., 2016; Sofroniew et al., 2016). In this respect, it aligns with current algorithms for cell detection in large-scale  $\text{Ca}^{2+}$  imaging data. Overall, our cell registration method expands the toolset available for analysis of time-lapse imaging data, offering the potential to explore coding dynamics within the same cells and neuronal ensembles over timescales of many weeks and even months.

## EXPERIMENTAL PROCEDURES

### Animals and $\text{Ca}^{2+}$ Imaging Procedures

All procedures were approved by the Weizmann Institute's Institutional Animal Care and Use Committee. For time-lapse imaging in freely behaving mice using an integrated miniature fluorescence microscope (nVistaHD, Insopix), we followed a previously established protocol (Ziv et al., 2013). We used data from a total of 12 male mice. Mice were housed in a reverse-light-cycle facility and were 8–12 weeks old at the beginning of the study. For  $\text{Ca}^{2+}$  imaging in the hippocampal CA1, we used five Thy1-GCaMP6f mice (Jackson Laboratory, #025393) (Dana et al., 2014) and five adeno-associated virus (AAV)-injected C57BL/6 wild-type mice (data from four out of five wild-type mice were taken from our previous work; Rubin et al., 2015). For  $\text{Ca}^{2+}$  imaging in the prefrontal cortex, we used two CaMKII-tTA and rtTA-GCaMP6s double-transgenic mice (CamKII-GCaMP6s, for short) (Tg(tetO-GCaMP6s)2Niell; Jackson Laboratory, #024742; and B6.Cg-Tg(CaMK2a-tTA)1Mmay/Dbo; Jackson Laboratory, #007004).  $\text{Ca}^{2+}$  imaging of the virus-injected and transgenic mice were performed at 20 Hz and 10 Hz, respectively. Additionally, we obtained published two-photon imaging data from ten different experiments from the Allen Brain Observatory, 2016 (<http://observatory.brain-map.org/visualcoding>). For a detailed description of the surgical and  $\text{Ca}^{2+}$  imaging procedures, see the Supplemental Experimental Procedures.

### Processing of $\text{Ca}^{2+}$ Imaging Data

We processed imaging data using commercial software (Mosaic, v1.1.1b, Insopix) and custom MATLAB routines, as previously described (Rubin et al., 2015). Briefly, the procedure consisted of the following main stages: (1) spatial down-sampling (final pixel size of  $2.3 \mu\text{m} \times 2.3 \mu\text{m}$ ) to increase computational speed; (2) image registration to correct for lateral displacements of the brain; (3) transformation into relative changes in fluorescence,  $\Delta F(t)/F_0 = (F(t) - F_0)/F_0$ , where  $F_0$  is the value for each pixel averaged over

time; (4) detection of cells and extraction of their spatial footprints using PCA-ICA (Mukamel et al., 2009); (5) extraction of each cell's fluorescence trace by applying the spatial footprints to the  $\Delta F(t)/F_0$  movies; and (6) detection of  $\text{Ca}^{2+}$  events using thresholding of the fluorescence trace. As an independent technique for cell detection, we also used CNMF-E (Zhou et al., 2016), an extension of the CNMF method (Pnevmatikakis et al., 2016), for one-photon microendoscopic data. For a detailed description of the processing routines, see the [Supplemental Experimental Procedures](#).

### Registration of Cells across Sessions

We developed a probabilistic method for registering cells across multiple sessions based on the similarity in their spatial footprints. We used two similarity measures, the distance between centers of mass of two spatial footprints (centroid distance) and the Pearson correlation between spatial footprints (spatial correlation), to decide whether pairs of neighboring cells from two different sessions are the same entity. We registered cells using three main stages: (1) image alignment; (2) probabilistic modeling of the data; and (3) cell registration with a clustering procedure.

#### (1) Image Alignment

For each session, we projected the centroid locations of all the cells' spatial footprints onto a single image. We computed the cross-correlations between the projections of each session with the projection of a reference session, examining different rotations. We corrected for translation and rotation differences between the sessions by applying the translations and rotations that resulted in the maximal cross-correlation, yielding each cell's location in the reference coordinate system.

#### (2) Probabilistic Modeling of the Data

We modeled the distributions of centroid distances (centroid distances model) and spatial correlations (spatial correlations model) as a weighted sum of the distributions of two subpopulations of cell-pairs, representing same cells and different cells. Specifically, we modeled the distribution of centroid distances for different cells as a multiplication of a sigmoid function by a linear function. To model the distribution of centroid distances for the same cells, we used a log-normal distribution to match the requirement of a non-negative distribution. Similarly, the distribution of spatial correlations for same cells was modeled as a log-normal distribution. Since the distribution of spatial correlations peaks close to 1, we defined the origin as 1 by taking 1 minus the spatial correlation value. We modeled the distribution of spatial correlations for different cells using a beta function to obtain a single peaked distribution between 0 and 1. For each model, we found the parameters of the distributions and weights that best fit the data in terms of the mean squared error (performed for each mouse individually). Applying similar principles, we modeled the joint distribution of centroid distances and spatial correlations (joint model) for neighboring cell-pairs. For a detailed description of the probabilistic models, see the [Supplemental Experimental Procedures](#) and Figure S6. To obtain the probability for any pair of neighboring cells from different sessions to be the same cell, given their spatial correlation and centroid distance, we calculated  $P_{\text{same}}(\text{dist}, \text{corr})$ , according to Bayes's rule:

$$P_{\text{same}}(\text{dist}, \text{corr}) = \frac{P(\text{dist}, \text{corr} | \text{same}) \cdot W_{\text{same}}}{P(\text{dist}, \text{corr} | \text{same}) \cdot W_{\text{same}} + P(\text{dist}, \text{corr} | \text{diff}) \cdot W_{\text{diff}}}$$

where  $W_{\text{same}}$  and  $W_{\text{diff}}$  are the estimated weights (prior probabilities) for the two subpopulations of same cells and different cells, respectively.  $P(\text{dist}, \text{corr} | \text{same})$  and  $P(\text{dist}, \text{corr} | \text{diff})$  are the estimated joint distributions of centroid distances and spatial correlations for same cells and different cells, respectively. Similar calculations were performed for each of the one-dimensional models.

#### (3) Cell Registration with a Clustering Procedure

We applied an iterative clustering procedure (Bansal et al., 2004) based on the estimated  $P_{\text{same}}$ . First, we created an initial list of cells based on a  $P_{\text{same}}$  threshold. The iterative clustering procedure was then applied, and each cell was clustered with the candidate from all the different sessions with which it had the highest  $P_{\text{same}}$ . If  $P_{\text{same}}$  was lower than the registration threshold, the cell formed a new cluster. The iterative process continued until the clustering procedure converged.

### Quantification of Registration Accuracy

#### Cell-Pairs

For the centroid distances model, the false-negative rate was estimated as the fraction of the model of same cells that is higher than the registration threshold:

$$\text{False negative rate} = \frac{\#\text{false negatives}}{\#\text{false negatives} + \#\text{true positives}}$$

The false-positive rate was estimated as the fraction of the model of different cells that is lower than the registration threshold:

$$\text{False positive rate} = \frac{\#\text{false positives}}{\#\text{false positives} + \#\text{true negatives}}$$

Additionally, we calculated the Gini coefficient  $G_1$  and the overall estimated registration accuracy (see the [Supplemental Experimental Procedures](#)). Similar calculations were performed for the spatial correlations and joint models.

#### Cell Registers

The register score was calculated for each registered cell as the total number of reliable cell-pairs out of the total number of pairwise combinations of cell-pairs from all sessions. Each pair was considered reliable if it met three criteria (see the [Supplemental Experimental Procedures](#)). Accordingly, we defined the register score as:

$$\text{register score} = \frac{1}{N(N-1)} \sum_{k=1}^{\bar{N}} \sum_{\substack{m=1 \\ m \neq k}}^N \delta(k, m)$$

where  $N$  is the total number of sessions,  $\bar{N}$  is the number of active sessions for a given cell register, and  $\delta(k, m)$  is 1 if a cell-pair in the cluster is reliable and 0 otherwise. To examine the scalability of the method for a large number of sessions, registration was performed simultaneously for multiple sessions with the clustering procedure, choosing random subsets of the 16 sessions comprising 4, 8, or 12 sessions, and for the entire set of 16 sessions.

### Validation of Cell Registration

#### Exclusivity and Transitivity Measures

To compare exclusivity and transitivity of the data with those expected by chance, we obtained shuffled data by measuring the centroid distances between cell-pairs from different sessions, where each session was taken from a different mouse (see the [Supplemental Experimental Procedures](#)).

#### Place-Field Stability

We analyzed mouse behavior videos using a custom MATLAB (Mathworks) routine and considered periods wherein the mouse ran  $>3$  cm/s. We divided each linear track into 24 bins (4 cm each) and computed the activity map (event rate per bin) for each neuron by dividing the smoothed (Gaussian kernel:  $\sigma = 1.5$  bins, size = 5 bins) map of  $\text{Ca}^{2+}$  event numbers by the smoothed map of occupancy (Rubin et al., 2015). We defined the place field as the peak value of the activity map. For each place field with  $\geq 5$  events for a given session, we computed the spatial information using the unsmoothed event-rate map of each cell, as previously described (Markus et al., 1994) (see the [Supplemental Experimental Procedures](#)). Cells with statistically significant spatial information were considered as place cells. We then calculated the place-field positional shift and place-field correlation (Pearson correlation) for cell-pairs from different sessions for which at least one cell in the pair was considered a place cell. To compare the place-field stability of the data with the stability expected by chance, we obtained shuffled data by randomizing the identity of cells across sessions.

#### Simulated Data

Simulated data were obtained by using representative spatial footprints of cells that were detected in our data and placing them in a random location in the FOV. The minimal centroid distance between the locations of neighboring cells was set to  $7 \mu\text{m} \pm 1 \mu\text{m}$  (normally distributed) to match the distribution of the observed data. In each session, a fraction of the cells was simulated as active, and their locations were set as the original centroid locations with a jitter (see

the [Supplemental Experimental Procedures](#)). The fraction of active cells and the radius of the jitter were chosen to match those of the observed data. To obtain simulated data with different noise levels, we systematically increased the radius of the centroid location jitter. Additionally, we simulated the effects of registration errors on measurements of the stability of the functional properties of cells. To test whether registration errors could obscure conclusions drawn from the data, we systematically changed effect sizes by changing the percentages of same cells that maintain their coding properties across sessions. Effect size was defined as the difference between the mean stability of the population and shuffled data, divided by the stability SD.

### Statistical Analysis

For related sample analysis, we performed repeated-measures ANOVA. Greenhouse-Geisser estimates of sphericity were used for adjustment of degrees of freedom, and Bonferroni correction was performed when conducting multiple comparisons.

### Code and Data Availability

The cell registration software, including a sample dataset that enables the generation of most of the figures presented in this paper, is deposited in a GitHub repository that can be accessed freely: <https://github.com/zivlab/CellReg>.

### SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and one movie and can be found with this article online at <https://doi.org/10.1016/j.celrep.2017.10.013>.

### AUTHOR CONTRIBUTIONS

L.S. developed the cell registration method, designed and performed analyses, and wrote and edited the paper. A.R. developed the cell registration method, designed analyses, and edited the paper. N.B.-E. performed experiments and analyses. N.G. performed experiments. N.S. performed experiments. O.P. performed analyses. Y.Z. supervised the project, designed analyses, and wrote and edited the paper.

### ACKNOWLEDGMENTS

Y.Z. is incumbent of the Daniel E. Koshland Sr. career development chair and supported by grants from the Abraham and Sonia Rochlin Foundation, the Hyman T. Milgrom Trust, the Minerva Foundation, the Israel Science Foundation (grant 2184/14), the European Research Council (ERC-StG 638644), and the FP7 Marie Curie actions(CIG 630852). We thank Aia Haruvit, Linor Balilit-Turgeman, Meytar Zemer, Yoav Rechavi, Hadas Baumer, Gloria Sheintuch, Ofer Yizhar, Rony Paz, Jerome Lecocq, James Fitzgerald, and Omri Barak for helpful advice and comments on the manuscript. We also thank the Allen Institute for sharing their *in vivo* datasets.

Received: May 2, 2017

Revised: August 14, 2017

Accepted: October 3, 2017

Published: October 24, 2017

### REFERENCES

- Bansal, N., Blum, A., and Chawla, S. (2004). Correlation clustering. *Mach. Learn.* **56**, 89–113.
- Barreto, R.P.J., Ko, T.H., Jung, J.C., Wang, T.J., Capps, G., Waters, A.C., Ziv, Y., Attardo, A., Recht, L., and Schnitzer, M.J. (2011). Time-lapse imaging of disease progression in deep brain areas using fluorescence microendoscopy. *Nat. Med.* **17**, 223–228.
- Berdyyeva, T., Otte, S., Aluisio, L., Ziv, Y., Burns, L.D., Dugovic, C., Yun, S., Ghosh, K.K., Schnitzer, M.J., Lovenberg, T., and Bonaventure, P. (2014). Zolpidem reduces hippocampal neuronal activity in freely behaving mice: a large scale calcium imaging study with miniaturized fluorescence microscope. *PLoS ONE* **9**, e112068.
- Burgess, C.R., Ramesh, R.N., Sugden, A.U., Levandowski, K.M., Minnig, M.A., Fenselau, H., Lowell, B.B., and Andermann, M.L. (2016). Hunger-dependent enhancement of food cue responses in mouse postrhinal cortex and lateral amygdala. *Neuron* **91**, 1154–1169.
- Cai, D.J., Aharoni, D., Shuman, T., Shobe, J., Biane, J., Song, W., Wei, B., Veshkini, M., La-Vu, M., Lou, J., et al. (2016). A shared neural ensemble links distinct contextual memories encoded close in time. *Nature* **534**, 115–118.
- Chen, T.W., Wardill, T.J., Sun, Y., Pulver, S.R., Renninger, S.L., Baohan, A., Schreiter, E.R., Kerr, R.A., Orger, M.B., Jayaraman, V., et al. (2013). Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295–300.
- Dana, H., Chen, T.W., Hu, A., Shields, B.C., Guo, C., Looger, L.L., Kim, D.S., and Svoboda, K. (2014). Thy1-GCaMP6 transgenic mice for neuronal population imaging *in vivo*. *PLoS ONE* **9**, e108697.
- Dombeck, D.A., Khabbaz, A.N., Collman, F., Adelman, T.L., and Tank, D.W. (2007). Imaging large-scale neural activity with cellular resolution in awake, mobile mice. *Neuron* **56**, 43–57.
- Dombeck, D.A., Harvey, C.D., Tian, L., Looger, L.L., and Tank, D.W. (2010). Functional imaging of hippocampal place cells at cellular resolution during virtual navigation. *Nat. Neurosci.* **13**, 1433–1440.
- Driscoll, L.N., Pettit, N.L., Minderer, M., Chettih, S.N., and Harvey, C.D. (2017). Dynamic reorganization of neuronal activity patterns in parietal cortex. *Cell* **170**, 986–999.e16.
- Ghosh, K.K., Burns, L.D., Cocker, E.D., Nimmerjahn, A., Ziv, Y., Gamal, A.E., and Schnitzer, M.J. (2011). Miniaturized integration of a fluorescence microscope. *Nat. Methods* **8**, 871–878.
- Grewe, B.F., Langer, D., Kasper, H., Kampa, B.M., and Helmchen, F. (2010). High-speed *in vivo* calcium imaging reveals neuronal network activity with near-millisecond precision. *Nat. Methods* **7**, 399–405.
- Grewe, B.F., Gründemann, J., Kitch, L.J., Lecoq, J.A., Parker, J.G., Marshall, J.D., Larkin, M.C., Jercog, P.E., Grenier, F., Li, J.Z., et al. (2017). Neural ensemble dynamics underlying a long-term associative memory. *Nature* **543**, 670–675.
- Grienberger, C., and Konnerth, A. (2012). Imaging calcium in neurons. *Neuron* **73**, 862–885.
- Harris, K.D., Quiroga, R.Q., Freeman, J., and Smith, S.L. (2016). Improving data quality in neuronal population recordings. *Nat. Neurosci.* **19**, 1165–1174.
- Helmchen, F., Fee, M.S., Tank, D.W., and Denk, W. (2001). A miniature head-mounted two-photon microscope. high-resolution brain imaging in freely moving animals. *Neuron* **31**, 903–912.
- Huber, D., Gutinsky, D.A., Peron, S., O'Connor, D.H., Wiegert, J.S., Tian, L., Oertner, T.G., Looger, L.L., and Svoboda, K. (2012). Multiple dynamic representations in the motor cortex during sensorimotor learning. *Nature* **484**, 473–478.
- Jennings, J.H., Ung, R.L., Resendez, S.L., Stamatakis, A.M., Taylor, J.G., Huang, J., Veleta, K., Kantak, P.A., Aita, M., Shilling-Scrivo, K., et al. (2015). Visualizing hypothalamic network dynamics for appetitive and consummatory behaviors. *Cell* **160**, 516–527.
- Kentros, C.G., Agnihotri, N.T., Streater, S., Hawkins, R.D., and Kandel, E.R. (2004). Increased attention to spatial context increases both place field stability and spatial memory. *Neuron* **42**, 283–295.
- Kim, T.H., Zhang, Y., Lecoq, J., Jung, J.C., Li, J., Zeng, H., Niell, C.M., and Schnitzer, M.J. (2016). Long-term optical access to an estimated one million neurons in the live mouse cortex. *Cell Rep.* **17**, 3385–3394.
- Kitamura, T., Sun, C., Martin, J., Kitch, L.J., Schnitzer, M.J., and Tonegawa, S. (2015). Entorhinal cortical ocean cells encode specific contexts and drive context-specific fear memory. *Neuron* **87**, 1317–1331.
- Kitamura, T., Ogawa, S.K., Roy, D.S., Okuyama, T., Morrissey, M.D., Smith, L.M., Redondo, R.L., and Tonegawa, S. (2017). Engrams and circuits crucial for systems consolidation of a memory. *Science* **356**, 73–78.

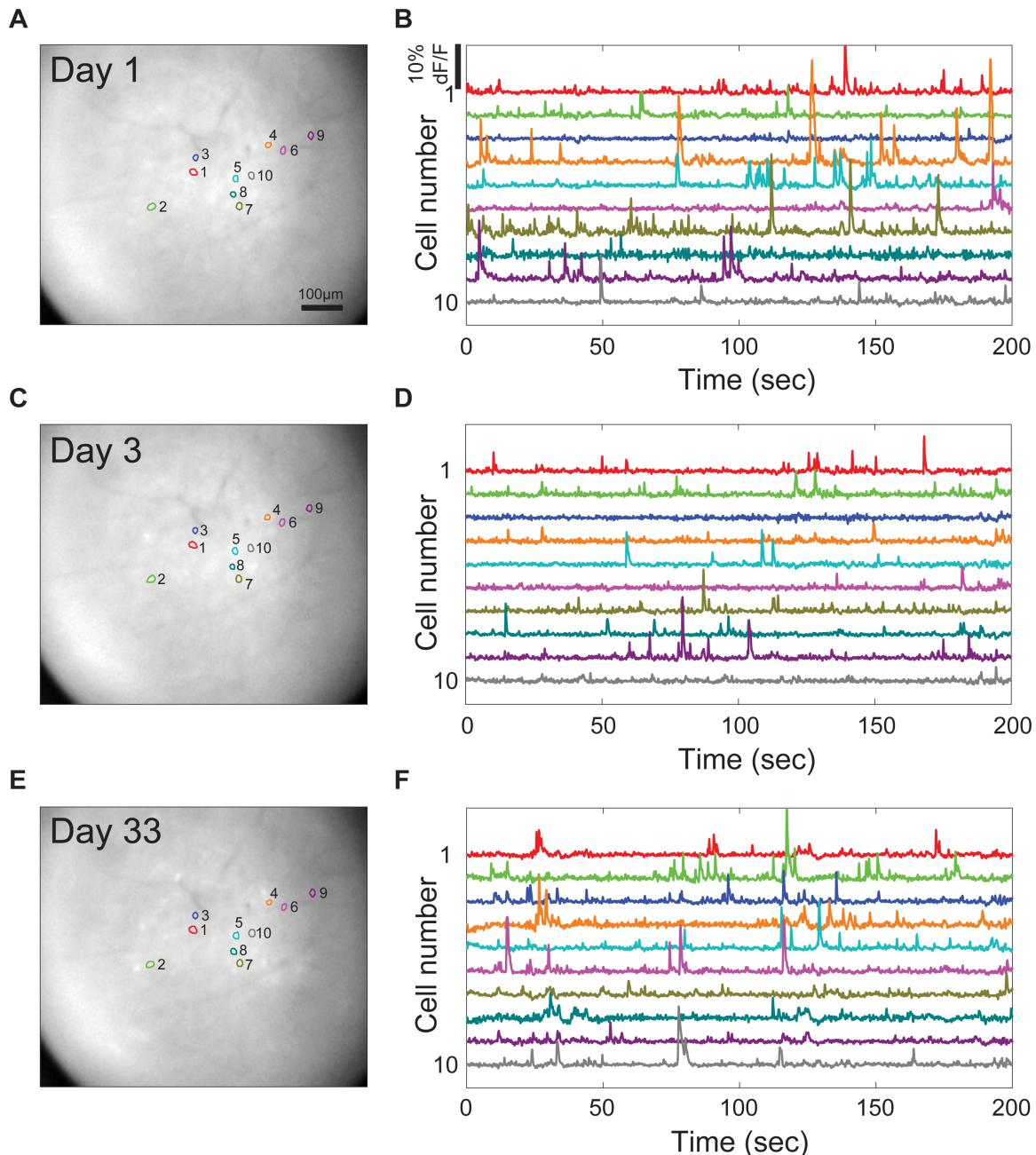
- Liberti, W.A.I., 3rd, Markowitz, J.E., Perkins, L.N., Liberti, D.C., Leman, D.P., Guitchounts, G., Velho, T., Kotton, D.N., Lois, C., and Gardner, T.J. (2016). Unstable neurons underlie a stable learned behavior. *Nat. Neurosci.* 19, 1665–1671.
- Lütcke, H., Margolis, D.J., and Helmchen, F. (2013). Steady or changing? Long-term monitoring of neuronal population activity. *Trends Neurosci.* 36, 375–384.
- Markus, E.J., Barnes, C.A., McNaughton, B.L., Gladden, V.L., and Skaggs, W.E. (1994). Spatial information content and reliability of hippocampal CA1 neurons: effects of visual input. *Hippocampus* 4, 410–421.
- Maruyama, R., Maeda, K., Moroda, H., Kato, I., Inoue, M., Miyakawa, H., and Aonishi, T. (2014). Detecting cells using non-negative matrix factorization on calcium imaging data. *Neural Netw.* 55, 11–19.
- Mukamel, E.A., Nimmerjahn, A., and Schnitzer, M.J. (2009). Automated analysis of cellular signals from large-scale calcium imaging data. *Neuron* 63, 747–760.
- O'Keefe, J., and Dostrovsky, J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Res.* 34, 171–175.
- Onativia, J., Schultz, S.R., and Dragotti, P.L. (2013). A finite rate of innovation algorithm for fast and accurate spike detection from two-photon calcium imaging. *J. Neural Eng.* 10, 046017.
- Pachitariu, M., Packer, A., Pettit, N., Dagleish, H., Hausser, M., and Sahani, M. (2013). Extracting regions of interest from biological images with convolutional sparse block coding. In *Advances in Neural Information Processing Systems 26 (NIPS 13)*, C.J.C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K.Q. Weinberger, eds. (Curran Associates), pp. 1745–1753.
- Pachitariu, M., Stringer, C., Schröder, S., Dipoppa, M., Rossi, L.F., Carandini, M., and Harris, K.D. (2016). Suite2p: beyond 10,000 neurons with standard two-photon microscopy. *bioRxiv*. <https://doi.org/10.1101/061507>.
- Pinto, L., and Dan, Y. (2015). Cell-type-specific activity in prefrontal cortex during goal-directed behavior. *Neuron* 87, 437–450.
- Pnevmatikakis, E.A., Soudry, D., Gao, Y., Machado, T.A., Merel, J., Pfau, D., Reardon, T., Mu, Y., Lacefield, C., Yang, W., et al. (2016). Simultaneous denoising, deconvolution, and demixing of calcium imaging data. *Neuron* 89, 285–299.
- Reidl, J., Starke, J., Omer, D.B., Grinvald, A., and Spors, H. (2007). Independent component analysis of high-resolution imaging data identifies distinct functional domains. *Neuroimage* 34, 94–108.
- Resendez, S.L., Jennings, J.H., Ung, R.L., Namboodiri, V.M., Zhou, Z.C., Otis, J.M., Nomura, H., McHenry, J.A., Kosyk, O., and Stuber, G.D. (2016). Visualization of cortical, subcortical and deep brain neural circuit dynamics during naturalistic mammalian behavior with head-mounted microscopes and chronically implanted lenses. *Nat. Protoc.* 11, 566–597.
- Rose, T., Jaepel, J., Hübener, M., and Bonhoeffer, T. (2016). Cell-specific restoration of stimulus preference after monocular deprivation in the visual cortex. *Science* 352, 1319–1322.
- Rubin, A., Geva, N., Sheintuch, L., and Ziv, Y. (2015). Hippocampal ensemble dynamics timestamp events in long-term memory. *eLife* 4, e12247.
- Sawinski, J., Wallace, D.J., Greenberg, D.S., Grossmann, S., Denk, W., and Kerr, J.N. (2009). Visually evoked activity in cortical cells imaged in freely moving animals. *Proc. Natl. Acad. Sci. USA* 106, 19557–19562.
- Smith, S.L., and Häusser, M. (2010). Parallel processing of visual space by neighboring neurons in mouse visual cortex. *Nat. Neurosci.* 13, 1144–1149.
- Sofroniew, N.J., Flickinger, D., King, J., and Svoboda, K. (2016). A large field of view two-photon mesoscope with subcellular resolution for *in vivo* imaging. *eLife* 5, e14472.
- Sun, C., Kitamura, T., Yamamoto, J., Martin, J., Pignatelli, M., Kitch, L.J., Schnitzer, M.J., and Tonegawa, S. (2015). Distinct speed dependence of entorhinal island and ocean cells, including respective grid cells. *Proc. Natl. Acad. Sci. USA* 112, 9466–9471.
- Svoboda, K., Denk, W., Kleinfeld, D., and Tank, D.W. (1997). *In vivo* dendritic calcium dynamics in neocortical pyramidal neurons. *Nature* 385, 161–165.
- Theis, L., Berens, P., Froudarakis, E., Reimer, J., Román Rosón, M., Baden, T., Euler, T., Tolias, A.S., and Bethge, M. (2016). Benchmarking spike rate inference in population calcium imaging. *Neuron* 90, 471–482.
- Thompson, L.T., and Best, P.J. (1990). Long-term stability of the place-field activity of single units recorded from the dorsal hippocampus of freely behaving rats. *Brain Res.* 509, 299–308.
- Vogelstein, J.T., Watson, B.O., Packer, A.M., Yuste, R., Jedynak, B., and Paninski, L. (2009). Spike inference from calcium imaging using sequential Monte Carlo methods. *Biophys. J.* 97, 636–655.
- Vogelstein, J.T., Packer, A.M., Machado, T.A., Sippy, T., Babadi, B., Yuste, R., and Paninski, L. (2010). Fast nonnegative deconvolution for spike train inference from population calcium imaging. *J. Neurophysiol.* 104, 3691–3704.
- Wagner, M.J., Kim, T.H., Savall, J., Schnitzer, M.J., and Luo, L. (2017). Cerebellar granule cells encode the expectation of reward. *Nature* 544, 96–100.
- Wilt, B.A., Burns, L.D., Wei Ho, E.T., Ghosh, K.K., Mukamel, E.A., and Schnitzer, M.J. (2009). Advances in light microscopy for neuroscience. *Annu. Rev. Neurosci.* 32, 435–506.
- Xia, L., Nygard, S.K., Sobczak, G.G., Hourguettes, N.J., and Bruchas, M.R. (2017). Dorsal-CA1 hippocampal neuronal ensembles encode nicotine-reward contextual associations. *Cell Rep.* 19, 2143–2156.
- Yang, W., and Yuste, R. (2017). *In vivo* imaging of neural activity. *Nat. Methods* 14, 349–359.
- Zhou, P., Resendez, S.L., Rodriguez-Romaguera, J., Jimenez, J.C., Neufeld, S.Q., Stuber, G.D., Hen, R., Kheirbek, M.A., Sabatini, B.L., Kass, R.E., et al. (2016). Efficient and accurate extraction of *in vivo* calcium signals from micro-endoscopic video data. *arXiv*, arXiv:1605.07266.
- Ziv, Y., and Ghosh, K.K. (2015). Miniature microscopes for large-scale imaging of neuronal activity in freely behaving rodents. *Curr. Opin. Neurobiol.* 32, 141–147.
- Ziv, Y., Burns, L.D., Cocker, E.D., Hamel, E.O., Ghosh, K.K., Kitch, L.J., El Gamal, A., and Schnitzer, M.J. (2013). Long-term dynamics of CA1 hippocampal place codes. *Nat. Neurosci.* 16, 264–266.

**Supplemental Information**

**Tracking the Same Neurons across Multiple Days  
in Ca<sup>2+</sup> Imaging Data**

**Liron Sheintuch, Alon Rubin, Noa Brande-Eilat, Nitzan Geva, Noa Sadeh, Or Pinchasof, and Yaniv Ziv**

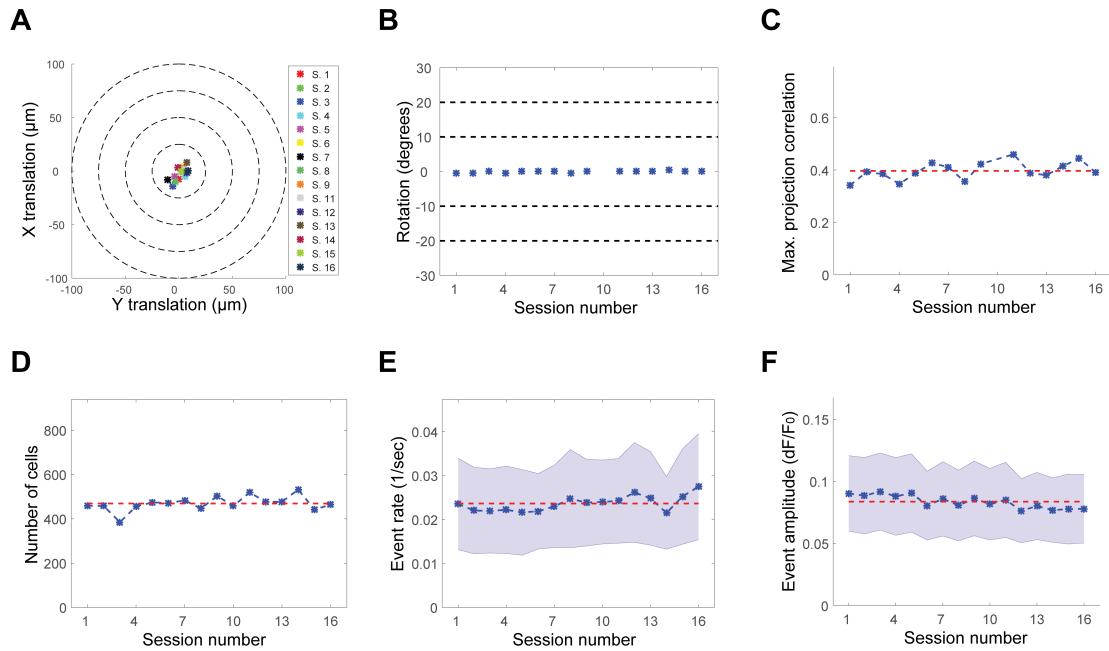
## Supplemental Figures and Legends



**Figure S1. The same cells are tracked over three sessions spanning more than a month.**

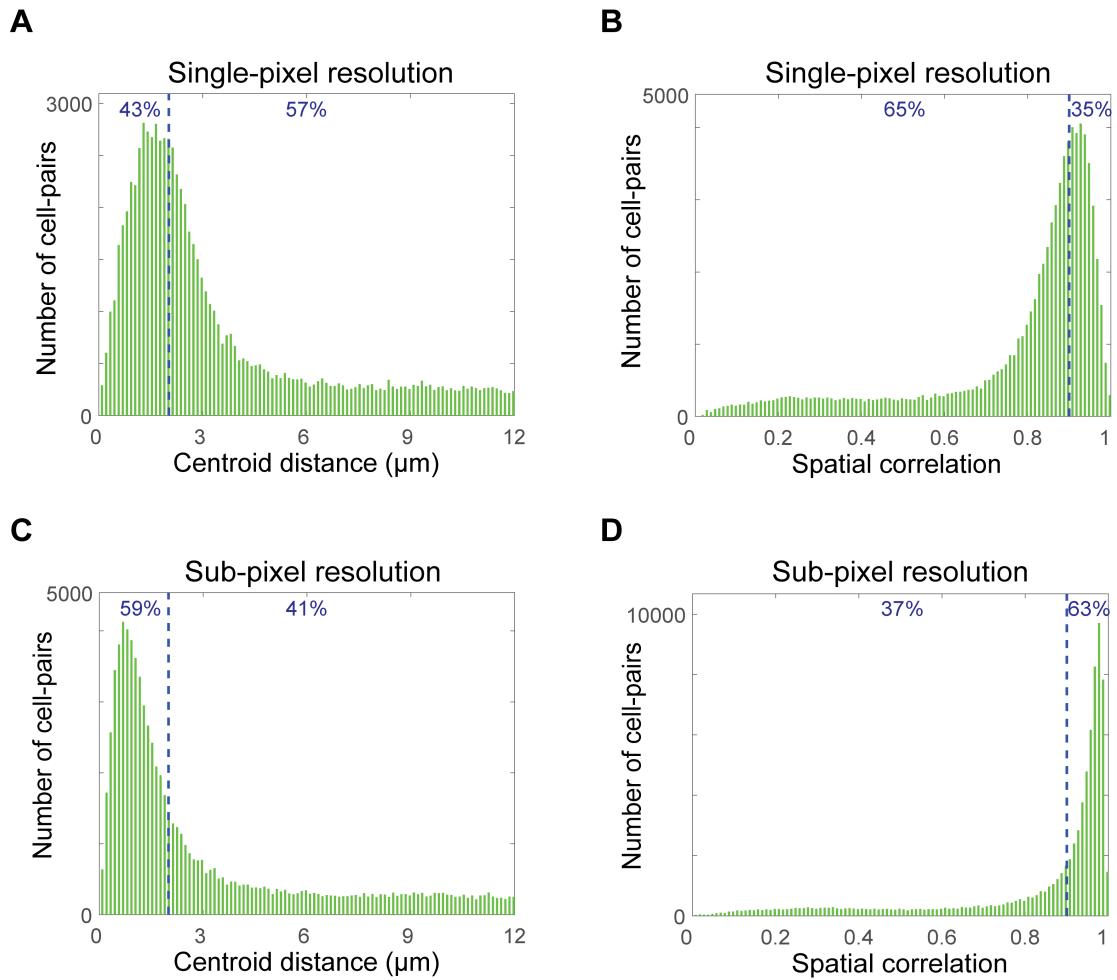
**Related to Figure 1.**

Identification of individual cells and extraction of their  $\text{Ca}^{2+}$  dynamics using principal component analysis and independent component analysis. (A,C,E) Contours of 10 spatial footprints overlaid on a representative single frame from day 1 (A), day 3 (C) and day 33 (E). (B,D,F) Corresponding  $\text{Ca}^{2+}$  traces for the same cells shown in A. Typical  $\text{Ca}^{2+}$  dynamics are observed for individual neurons. Data taken from three sessions recorded in the hippocampal CA1 of a Thy1-GCaMP6f transgenic mouse while freely exploring the same environment.



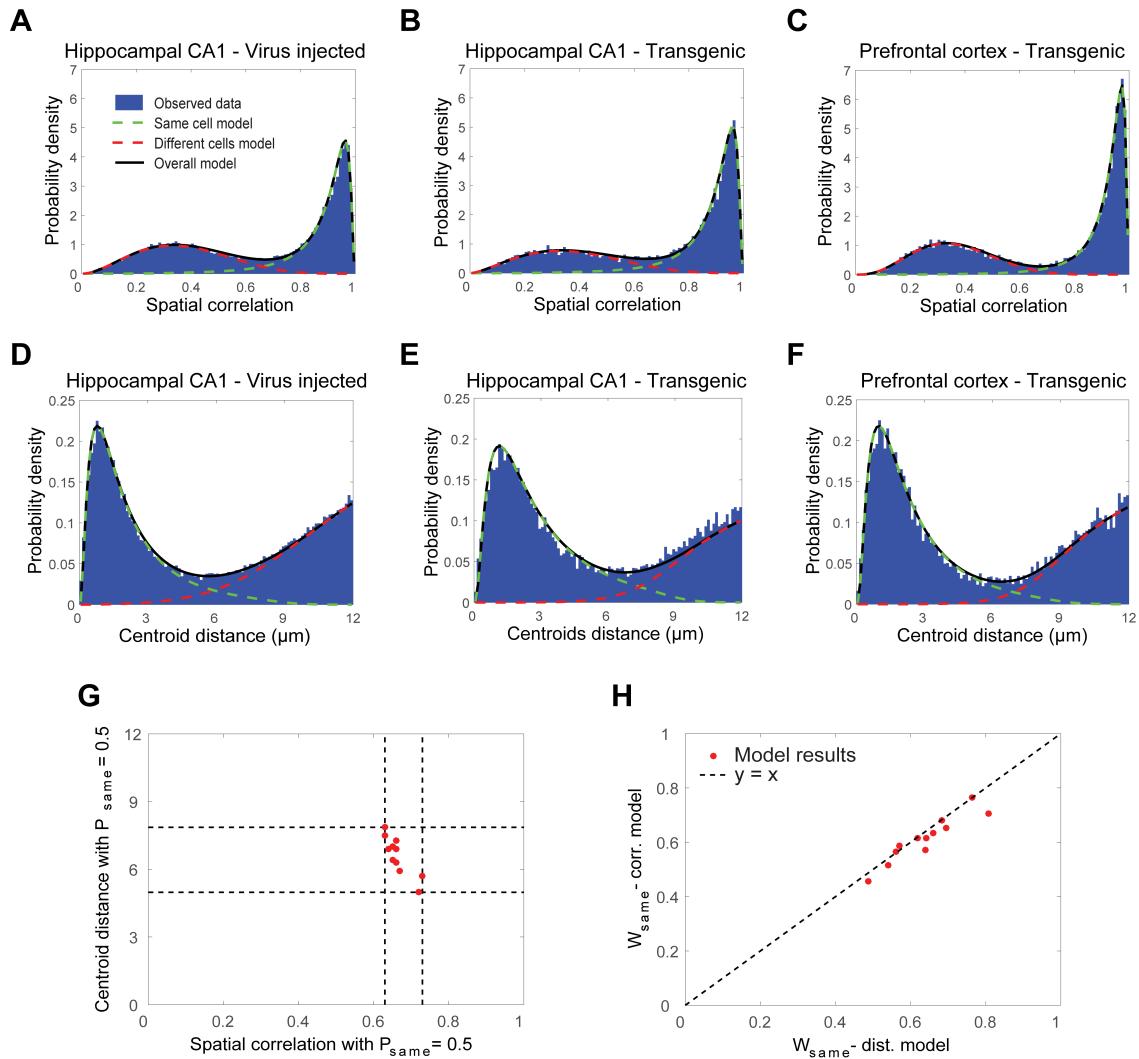
**Figure S2. A stable preparation is maintained across all days of the experiment. Related to Figure 1.**

(A,B) The translations (A) and rotations (B) that result in maximal correlation between the projection of the spatial footprint centroids in each session and that of the reference session. Session ten was chosen as a reference. (C) The maximal correlation between the projections of the spatial footprint centroids in each session and that of the reference session (blue asterisks), and the average maximal correlation across all sessions (red dashed curve). (D) The number of detected cells in each imaging session (blue asterisks) and the average number of cells across all sessions (red dashed curve). (E) The average cell event rate in each imaging session (SD is shown as shaded blue area), and the average cell event rate across all sessions (red dashed curve). (F) The average cell event amplitude in each imaging session (SD is shown as shaded blue area), and the average cell event amplitude across all sessions (red dashed curve). A-C can potentially point out deviations in the FOV of each session compared to the reference session, while significant deviations in the cellular activity detected in each session compared to the average cellular activity could potentially be uncovered in D-F. Data in all panels was taken from 16 sessions recorded on eight different days in the hippocampal CA1 of a mouse while freely exploring the same environments.



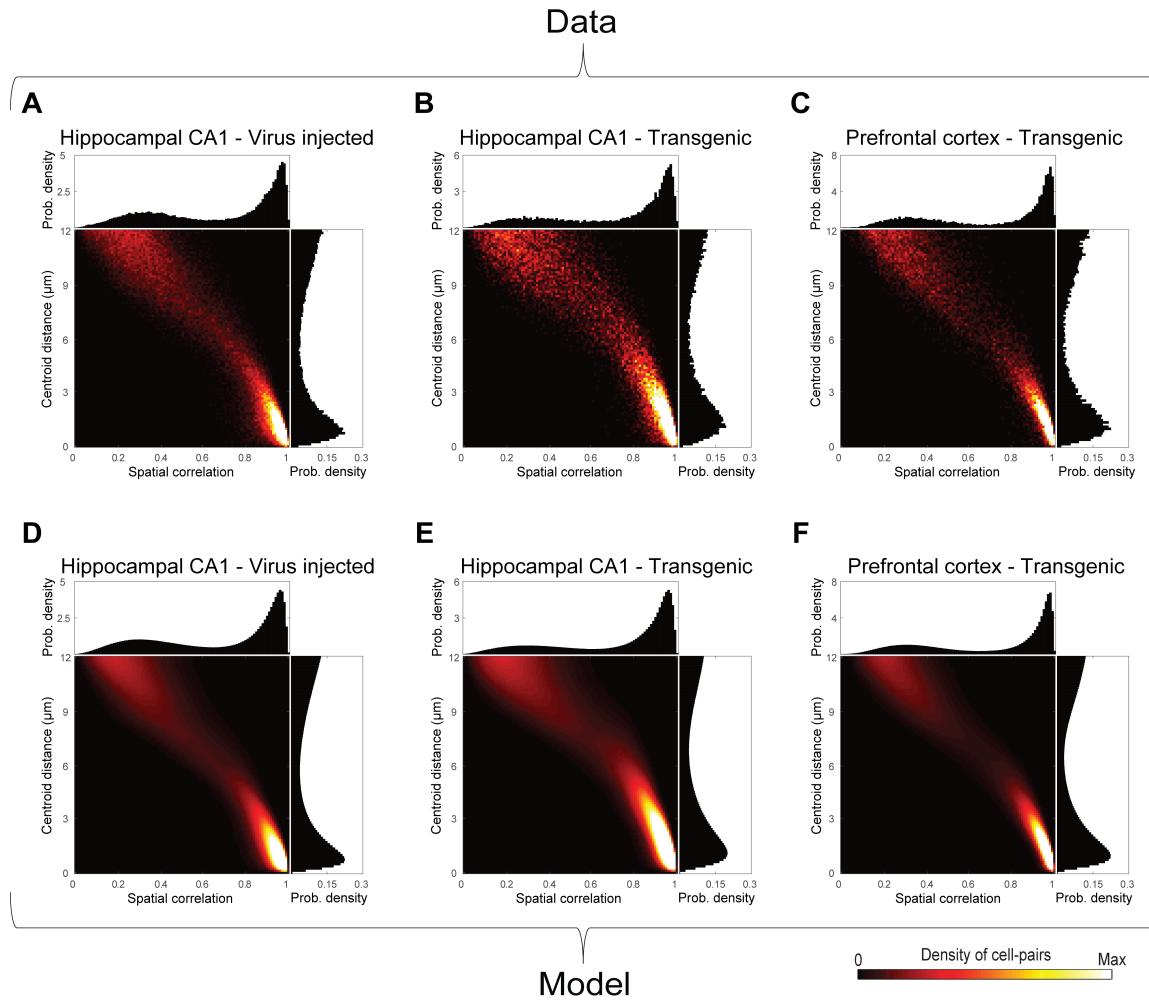
**Figure S3. Image alignment with sub-pixel resolution increases the spatial footprint similarity between pairs of nearest neighbor cells. Related to Figure 2.**

(A-D) Distributions of centroid distances and spatial correlations between pairs of nearest neighbor cells across days obtained after alignment (rigid-body registration) with single-pixel (A and B) and sub-pixel (C and D) resolution. We limited the distribution to cell-pairs with centroid distances  $<12\mu\text{m}$ . The blue dashed lines show centroid distances = $2\mu\text{m}$  and spatial correlation =0.9. Note the higher probabilities for nearest neighbors to have centroid distances  $<2\mu\text{m}$  and spatial correlation  $>0.9$  for alignment with sub-pixel resolution in comparison with their single-pixel counterparts. This fine spatial resolution of alignment was attainable because of the large number of centroid locations that were used to align each session. The data presented here were pre-processed with a pixel size of  $2.3 \times 2.3\mu\text{m}$  which is not negligible compared to the typical cell size. The effect of sub-pixel resolution on alignment should be less significant for imaging data with higher spatial resolution. Data in all panels was taken from 16 sessions recorded on eight different days in the hippocampal CA1 of a mouse while freely exploring the same environments.



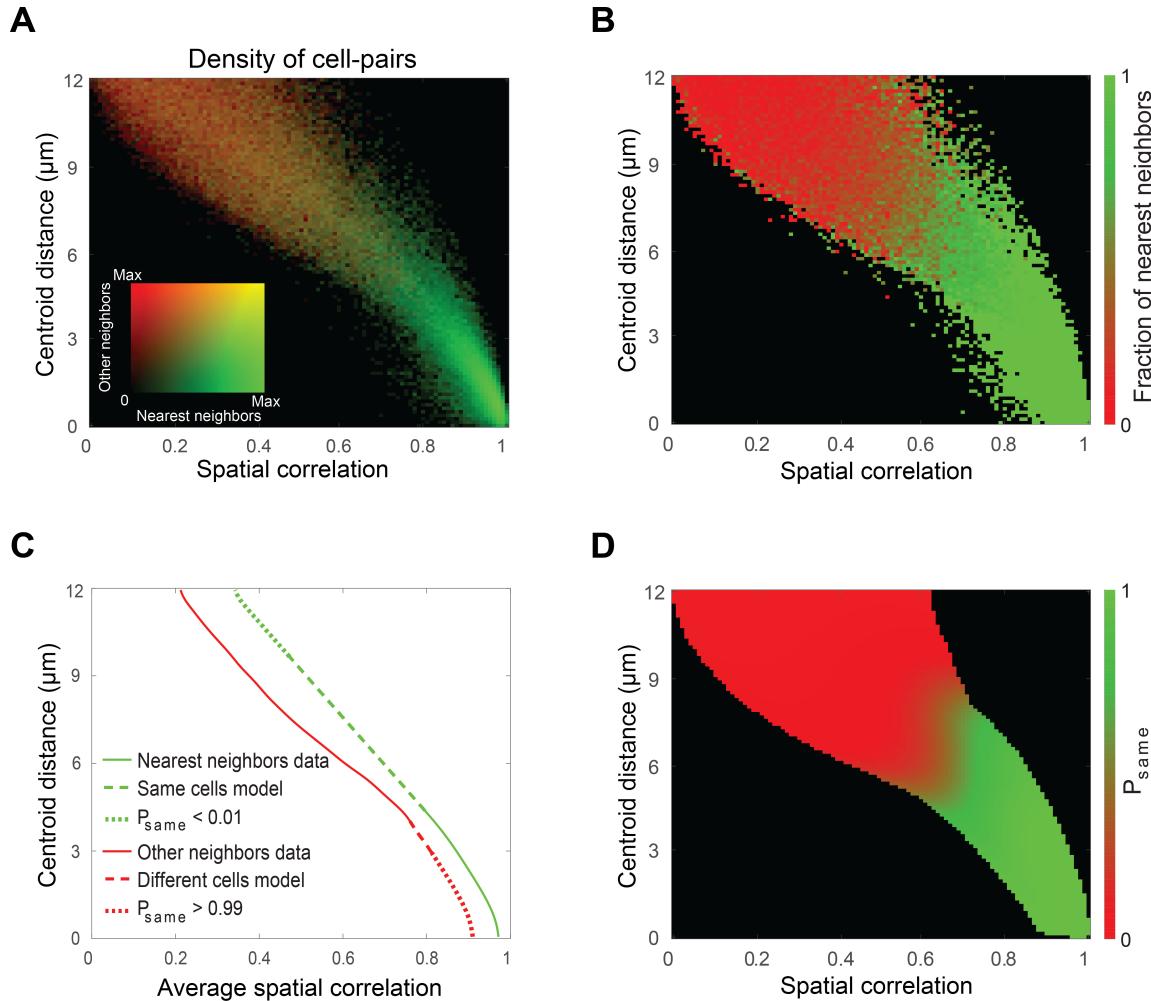
**Figure S4. Examples of the distributions of spatial correlations and centroid distances modeled for each mouse individual. Related to Figure 2.**

(A-F) Distributions of spatial correlations (A-C) and centroid distances (D-F) in three different mice for neighboring cell-pairs (blue bars), and the modeled distributions of same cells (dashed green curves), different cells (dashed red curves), and their weighted sum (solid black curves). Estimated distributions were obtained by assuming that the data consists of a weighted sum of two subpopulations, corresponding to same cells and different cells, and finding the parameters that best fit the data. Data and models in A and D are for 16 sessions recorded on eight different days in the hippocampal CA1 of a GCaMP6f virus injected mouse while freely exploring the same environments. Data and models in B and E are for seven sessions recorded on seven different days in the hippocampal CA1 of a Thy1-GCaMP6f transgenic mouse while freely exploring the same environment. Data and models in C and F are for six sessions recorded on six different days in the prefrontal cortex of a CaMKII-GCaMP6s transgenic mouse while freely exploring the same environments. (G) The intersection between the estimated distribution of same cells and different cells (corresponding to a registration threshold of  $P_{\text{same}}=0.5$ ) for the centroid distances model versus the spatial correlations model. Each dot represents a different mouse ( $N=12$ ). Black dashed lines show the minimum and maximum of these values across all mice. Note that data from different mice yielded different registration thresholds. (H) The weight of the subpopulation of same cells ( $W_{\text{same}}$ ) obtained by the centroid distances (dist.) model versus  $W_{\text{same}}$  obtained by the spatial correlations (corr.) model. Each dot represents a different mouse ( $N=12$ ). Black dashed line represents the  $y=x$  curve. Note that the spatial correlations and centroid distances models independently reach similar  $W_{\text{same}}$  values.



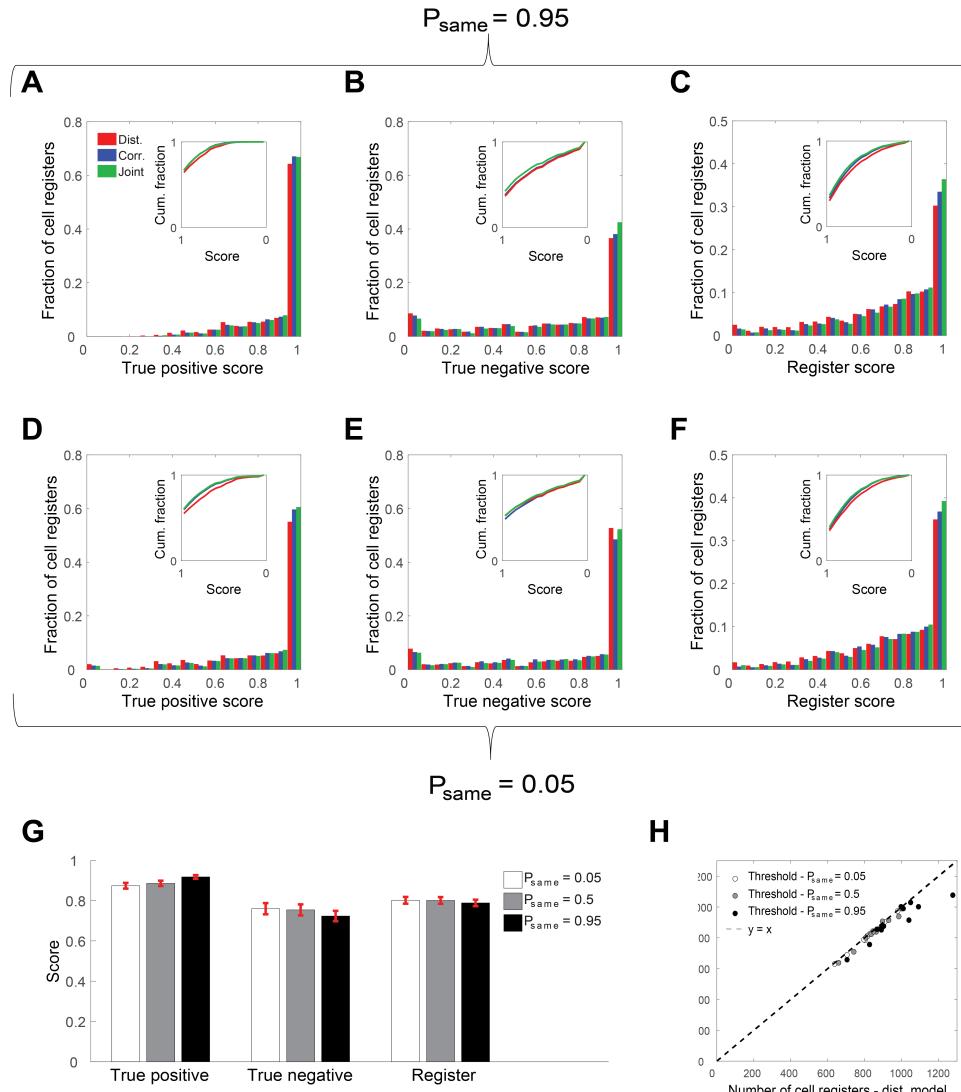
**Figure S5.** Examples of the joint distribution of spatial correlations and centroid distances modeled for each mouse individually. Related to Figure 2.

Joint distribution of spatial correlations and centroid distances and their corresponding marginal distributions for neighboring cell-pairs as observed in the data of three different mice (A-C), and the weighted sum of the models for same cells and different cells (D-F). The color scale was set to reach 0.25 of the maximal value to enable visualization of both subpopulations. Data and models in A and D are for 16 sessions recorded on 8 different days in the hippocampal CA1 of a GCaMP6f virus injected mouse while freely exploring the same environments. Data and models in B and E are for seven sessions recorded on seven different days in the hippocampal CA1 of a Thy1-GCaMP6f transgenic mouse while freely exploring the same environment. Data and models in C and F are for six sessions recorded on six different days in the prefrontal cortex of a CaMKII-GCaMP6s transgenic mouse while freely exploring the same environments.



**Figure S6. Obtaining the joint model for the joint distribution of spatial correlations and centroid distances based on nearest neighbors and other neighbors. Related to Figure 2.**

(A) An RG overlay of the density of nearest neighbor (green) and other neighbor (red) cell-pairs given the centroid distance and spatial correlation. (B) Ratio of the density of nearest neighbor to all neighbor cell-pairs given the centroid distance and spatial correlation. Note the high fraction of nearest neighbor cell-pairs for high spatial correlations and low centroid distances. (C) The average spatial correlation as a function of the centroid distance for nearest neighbors (solid green curve), and other neighbors (solid red curve), and extrapolations of the data (dashed and dotted green and red curves), used to obtain the models for same cells and different cells, respectively. Dotted green curve represents centroid distances for which  $P_{\text{same}} < 0.01$ , a region for which the extrapolation for the same cells model is insignificant, while the dotted red curve represents centroid distances for which  $P_{\text{same}} > 0.99$ , a region for which the extrapolation for the different cells model is insignificant. (D) The probability for two cells from two different sessions to be the same cell ( $P_{\text{same}}$ ) given the centroid distance and spatial correlation, as estimated by the model.  $P_{\text{same}}$  is presented exclusively for centroid distances and spatial correlations where the model's probability for cell-pairs of either type (same cell or different cells) is  $> 10^{-5}$ . Data and model in all panels are for 16 sessions recorded on eight different days from GCaMP6f-expressing CA1 neurons of a mouse while freely exploring the same environments.



**Figure S7. False positive and false negative rates can be controlled by the registration threshold. Related to Figure 5.**

True positive (A and D), true negative (B and E), and register (C and F) scores for cell registers with thresholds of  $P_{\text{same}}=0.95$  (A-C) and  $P_{\text{same}}=0.05$  (D-F), for the centroids distances (red), spatial correlations (blue), and joint (green) models. Insets, cumulative fraction of cell registers as a function of the score reversed from 1 to 0. (G) Average true positive, true negative and register scores for the joint model with thresholds of  $P_{\text{same}}=0.05$  (white bars),  $P_{\text{same}}=0.5$  (gray bars), and  $P_{\text{same}}=0.95$  (black bars). Error bars show SEM. The average true positive score increases while the average true negative score decreases with the registration threshold. Note, that using a more relaxed registration threshold ( $P_{\text{same}}=0.05$ ) results in a slightly higher average register score than a stringent threshold ( $P_{\text{same}}=0.95$ ). This result is partially explained by the fact that if more than one candidate exceeds the registration threshold, only the one with highest  $P_{\text{same}}$  will be registered as the same cell, thus eliminating a large fraction of false positives. (H) Final number of cell registers obtained by centroid distances based registration versus spatial correlations based registration. Dots represent the thresholds  $P_{\text{same}}=0.05$  (white circles),  $P_{\text{same}}=0.5$  (gray circles), and  $P_{\text{same}}=0.95$  (black circles) for each mouse. Black dashed line represents the  $y=x$  curve. Note that the spatial correlations and centroid distances models obtain similar numbers of cell registers and that increasing the registration threshold increases the final number of cell registers. Data in all panels are pooled from N=12 mice.

## **Supplemental Experimental Procedures**

### **Animals and surgical procedures**

All procedures were approved by the Weizmann's Institute Institutional Animal Care and Use Committee. We used data from a total of 12 male mice. Mice were housed in cages with running wheels in a reverse light cycle facility, and were 8-12 weeks old at the beginning of the study. For  $\text{Ca}^{2+}$  imaging in the hippocampal CA1, we used five Thy1-GCaMP6f mice (Jackson Laboratory, #025393) (Dana *et al.*, 2014), and five AAV injected C57BL/6 wild type mice (data from four out of five wild type mice were taken from our previous work ([Rubin et al., 2015](#))). For  $\text{Ca}^{2+}$  imaging in the prefrontal cortex, we used two CaMKII-tTA and rtTA-GCaMP6s double transgenic mice (CamKII-GCaMP6s, for short) (Tg(tetO-GCaMP6s)2Niell; Jackson Laboratory, #024742; B6.Cg-Tg(CaMK2a-tTA)1Mmay/DboJ; Jackson Laboratory, #007004; all mice were bred on C57BL/6 background). C57BL/6 wild type and Thy1-GCaMP6f mice were housed with 1-4 cage-mates. CamKII-GCaMP6s mice were single housed. All surgical procedures were conducted under isoflurane anesthesia (1.5-2% volume). We used a viral vector to express GCaMP6s or GCaMP6f in the CA1 of wild type mice. These mice underwent two surgical procedures. In the first, we injected into the CA1 400 nL of the viral vector AAV2/5-CaMKIIa-GCaMP6s or AAV2/5-CaMKIIa-GCaMP6f ([Chen et al., 2013](#)) (~2 X  $10^{13}$  particles per ml, packed by University of North Carolina Vector Core). Stereotactic coordinates were: -1.9 mm antero-posterior, -1.4 mm mediolateral, -1.6 mm dorsoventral from bregma. Mice were allowed to recover in their home-cages for at least one week before the next surgical procedure, in which we implanted a glass guide tube directly above the CA1, as previously described ([Ziv et al., 2013](#); [Rubin et al., 2015](#)). This procedure was similar for both virus injected and Thy1-GCaMP6f transgenic mice. CamKII-GCaMP6s mice were implanted with a micro-prism lens (800 $\mu\text{m}$  diameter) in the prefrontal cortex. Stereotactic coordinates of the implantation were: 1 mm anterior-posterior, 0 mm mediolateral, -1.8 mm dorsoventral from bregma.

### **$\text{Ca}^{2+}$ imaging and behavioral setup**

#### *Preparatory process*

For time-lapse imaging in freely behaving mice using an integrated miniature fluorescence microscope (nVistaHD, Inscopix), we followed a previously established protocol ([Ziv et al., 2013](#); [Rubin et al., 2015](#)). Briefly, at least three weeks after the guide tube or micro-prism lens implantation, we examined  $\text{Ca}^{2+}$  indicator expression and tissue health by imaging mice under isoflurane anesthesia using a two-photon microscope (Ultima IV, Bruker, Germany), equipped with a tunable Ti: Sapphire laser (Insight, Spectra Physics, Santa Clara, CA). At

this stage, we inserted into the guide tube of CA1 implanted mice a ‘microendoscope’ consisting of a single gradient refractive index lens (0.44 pitch length, 0.47 NA, GRINtech GmbH, Germany). We selected for further imaging only those mice that exhibited homogenous GCaMP6 expression and healthy appearance of the tissue. For the selected CA1 implanted mice, we affixed the microendoscope within the guide tube using ultraviolet-curing adhesive (Norland, NOA81, Edmund Optics, Barrington, NJ). Next, we attached the microscope’s base plate to the dental acrylic cap using light cured acrylic (Flow-It ALC, Pentron, Orange, CA). We used a similar procedure to attach the base plate to the skull of the mice that were implanted with a micro-prism lens. All mice were returned to their home cages for a few days following the aforementioned procedure.

#### *Ca<sup>2+</sup> imaging in freely behaving mice*

We used data from three experiments. In the first, we trained the five wild type mice to run back and forth on two elevated 96cm straight or ‘L’ shaped linear tracks ([Rubin et al., 2015](#)). Before beginning with Ca<sup>2+</sup> imaging, we trained the mice for 8–11 days, until the mice ran at least 60 times the entire length of each track in two consecutive days. We imaged the mice every other day for 15 days, making for eight recording days. Each day of the experiment consisted of two 15 minutes long sessions, separated by 4–5 hours. In the second experiment, we imaged the five Thy1-GCaMP6f mice while they were freely exploring a spatial environment (a total of seven 20-55 minutes long sessions, separated by 1-28 days). In the third experiment, we imaged the prefrontal cortex of the two CamKII-GCaMP6s mice on six consecutive days, while the mice were freely exploring two different environments (3.5 minutes per day in each environment). Ca<sup>2+</sup> imaging of the virus injected and transgenic mice was performed at 20Hz and 10Hz, respectively. To record mouse behavior, we used an overhead camera (DFK 33G445, The Imaging Source, Germany), which we synchronized with the integrated microscope.

#### **Two-photon imaging data**

We obtained published two-photon imaging data from the [Allen Brain Observatory, 2016](#) (<http://observatory.brain-map.org/visualcoding>). We used data from ten different experiments in head fixed mice (experiment numbers: 511510945; 511510779; 511510699; 511510670; 511510911; 511510718; 511510870; 511510836; 511856567; 511498742). The experiments measured visual responses from GCaMP6f expressing neurons in different cortical areas and layers. Each experiment consisted of three recording sessions from the same mouse, separated by at least one day.

### Processing of $\text{Ca}^{2+}$ imaging data

We processed imaging data using commercial software (Mosaic, version 1.1.1b, Inscopix) and custom MATLAB routines as previously described ([Ziv et al., 2013](#); [Rubin et al., 2015](#)). To increase computation speed, we spatially down-sampled the data by a factor of two in each dimension (final pixel size of  $2.3 \times 2.3\mu\text{m}$ ). To correct for non-uniform illumination both in space and time, we normalized the images by dividing each pixel by the corresponding value of that pixel in a smoothed version. The smoothed version was obtained by applying a Gaussian filter with a radius of  $100\mu\text{m}$  on the movies. Normalization also enhanced the appearance of the blood vessels, which were later used as stationary fiducial markers for image registration. We used a rigid-body registration to correct for lateral displacements of the brain. This procedure was performed on a high contrast subregion of the normalized movies for which the blood vessels were most prominent. The movies were then transformed into relative changes in fluorescence,  $\Delta F(t)/F_0 = (F(t) - F_0)/F_0$ , where  $F_0$  is the value for each pixel averaged over time. For the purpose of cell detection, the movies were down-sampled in time by a factor of five or two for data recorded at 20Hz and 10Hz, respectively. We detected *spatial footprints* (i.e., weighted regions of interest consisting of each pixel's contribution to the cell's fluorescence) corresponding to individual cells using an established cell detection algorithm that applies principal and independent component analyses (PCA and ICA; ([Mukamel, Nimmerjahn and Schnitzer, 2009](#))). For each spatial footprint, we used a threshold of 50% of the footprint's maximum intensity, and each pixel that did not cross the threshold was set to zero. Once the cells were detected, further cell sorting was performed to find the spatial footprints that follow a typical cellular structure. This was done by measuring the footprints' area and circularity and discarding those whose radius was smaller than  $5\mu\text{m}$  or larger than  $15\mu\text{m}$ , or which had a circularity smaller than 0.8. In some cases, the output of the PCA-ICA algorithm included more than one component that corresponded to a single cell. To eliminate such occurrences, we examined all pairs of cells with centroid distances  $<18\mu\text{m}$ , and whenever their traces had a correlation  $>0.9$ , the cell with the lower average event peak amplitude was discarded. As an independent technique for cell detection, we also used CNMF-E ([Zhou et al., 2016](#)), an extension of the constrained nonnegative matrix factorization method ([Pnevmatikakis et al., 2016](#)) for one-photon microendoscopic data.

### Detection of $\text{Ca}^{2+}$ events

$\text{Ca}^{2+}$  activity was extracted by applying the thresholded spatial footprints to the full temporal resolution (20Hz or 10Hz)  $\Delta F(t)/F_0$  movies. Baseline fluctuations were removed by subtracting the median trace (20sec sliding window). The  $\text{Ca}^{2+}$  traces were smoothed with a

low-pass filter with a cutoff frequency of 2Hz.  $\text{Ca}^{2+}$  candidate events were detected whenever the amplitude crossed a threshold of 4 or 5 median absolute deviations (MAD), for GCaMP6s or GCaMP6f, respectively. We considered for further analysis only candidate  $\text{Ca}^{2+}$  events with decay time equal or longer than 600msec or 200msec for GCaMP6s or GCaMP6f, respectively, consistent with typically observed indicator decay times (Chen et al., 2013). To avoid the detection of several peaks for a single  $\text{Ca}^{2+}$  event, only peaks that were 4 or 5 MAD higher than the previous peak (within the same candidate event) and 2 or 2.5 MAD higher than the next peak for GCaMP6s or GCaMP6f, respectively, were regarded as true events. We set the  $\text{Ca}^{2+}$  event occurrence to the time of the peak fluorescence. To mitigate the effects of crosstalk (i.e., spillover of  $\text{Ca}^{2+}$  fluorescence from neighboring cells), we adopted a conservative approach, allowing only one cell from a group of neighbors ( $<18\mu\text{m}$  apart) to register a  $\text{Ca}^{2+}$  event in a 200msec time window (the event with highest peak  $\Delta F(t)/F_0$  value). If two neighboring cells had a correlation  $>0.9$  in their events, the cell with the lower average peak amplitude was discarded. Once the events were identified, further event sorting was performed to find the cells with sufficient signal-to-noise ratios. This was accomplished by measuring the event rate and the average event peak amplitude for each cell and discarding those whose event rate was smaller than 0.01Hz or which had an average event amplitude smaller than 1%  $\Delta F(t)/F_0$ .

### Registration of cells across sessions

We developed a probabilistic method for registering cells across multiple sessions based on the similarity in the spatial footprint of their cellular activity. We used two similarity measures, the distance between centers of mass of two spatial footprints (*centroid distance*) and the Pearson correlation between spatial footprints (*spatial correlation*), to decide whether or not pairs of neighboring cells from two different sessions are the same entity. Note, that the spatial correlation is not merely a measure of shape similarity (as in spatial cross-correlation), but also accounts for differences in the cells' locations within the FOV. We registered cells using three main stages: (1) image alignment: rigid-body transformation was applied to all sessions according to a single reference session to correct for translation and rotation differences between sessions; (2) the distribution of spatial footprint similarities between neighboring cells from different sessions was calculated and then modeled as a weighted sum of the distributions of two subpopulations representing the same cells and different cells. This allowed us to estimate the probability for any pair of neighboring cells from two different sessions to be the same cell ( $P_{\text{same}}$ ) given their spatial correlation and

centroid distance; (3) an iterative clustering procedure ([Bansal, Blum and Chawla, 2004](#)) that registers cells based on the  $P_{\text{same}}$  obtained from the probabilistic model.

### *(1) Image alignment*

For each session, we projected the centroid locations of all the cells' spatial footprints onto a single image. We computed the cross-correlations between the projections of each session with the projections of a reference session, examining different rotations. We corrected for translation and rotation differences between the sessions by applying the translations and rotations that resulted in the maximal cross-correlation, yielding each cell's location in the reference coordinate system. We also corrected for sub-pixel translational differences between the sessions by applying a Gaussian fit to the cross-correlation in each dimension. Sub-pixel translational correction decreased the centroid distances and increased the spatial correlations between the nearest neighbors ([Figure S3](#)).

### *(2) Probabilistic modeling of the data*

We calculated the distributions of centroid distances and spatial correlations for neighboring cell-pairs across sessions. We defined neighboring cell-pairs as pairs with centroid distances  $<12\mu\text{m}$  across sessions, and assumed that cell-pairs with larger centroid distances are different cells. To obtain reliable registration we sought to model the distributions of centroid distances (centroid distances model) and spatial correlations (spatial correlations model) for neighboring cell-pairs, as a weighted sum of the distributions of two subpopulations, representing same cells and different cells. The specific choice of the functional forms we used to model the distributions of same cells and different cells were guided by the distributions obtained from the data for nearest neighbors and other neighbors.

Because there is a certain minimal distance between different cells in the FOV, we approximated the density of cells given their distance from any given cell as a sigmoid function. Because the area of a ring surrounding each cell linearly increases with the distance, we modeled the distribution of centroid distances for different cells as a multiplication of a sigmoid function by a linear function. To model the distribution of centroid distances for the same cells, we used a lognormal distribution, to match the requirement of a non-negative distribution that resembles the distribution for nearest neighbors in the data. Similarly, the distribution of spatial correlations for same cells was modeled as a lognormal distribution. Since the correlation cannot exceed 1, and the distribution of spatial correlations for nearest neighbors peaks close to 1 and decays towards 0, we defined the origin as 1 by taking 1 minus the spatial correlation value. Because lognormal distribution is  $>0$  for any positive value, whereas correlation is bounded, we multiplied the obtained model for same

cells by a sigmoid function. The distribution of spatial correlations for other neighbors is single peaked and ranges between 0 and 1. We therefore modeled the distribution of spatial correlations for different cells using a beta function. For each model, we found the parameters of the distributions and weights that best fit the data in terms of the mean squared error (performed for each mouse, individually).

Applying similar principles, we modeled the joint distribution of centroid distances and spatial correlations (joint model) for neighboring cell-pairs as a weighted sum of the distributions of two subpopulations. We used the previously estimated distributions of centroid distances and their weights as the marginal distribution of the joint model. We then estimated the distribution of spatial correlations for any given centroid distance ([Figure S6](#)). We calculated the average spatial correlation for nearest neighbors and other neighbors given the centroid distance to approximate the spatial correlation for same cells and different cells, respectively. We used the average spatial correlation measured for nearest neighbors at small centroid distances (mostly the same cells) to linearly extrapolate the average spatial correlation for large centroid distances where we could not be certain that the nearest neighbors were indeed the same cells. Assuming that other neighbors correspond to different cells, we used the average spatial correlation measured for other neighbors at high centroid distances, and extrapolated the data to small centroid distances for which data was sparse. For a given centroid distance, the nearest neighbors showed higher spatial correlations than other neighbors. We therefore measured the difference between the average spatial correlation for nearest neighbors and other neighbors per centroid distance and extrapolated the mean spatial correlation for different cells assuming a constant difference. In practice, this extrapolation had little impact on the joint model since it was calculated mostly for a centroid distances range in which the proportion of different cells was < 0.01 ([Figure S6C](#) – note the narrow range of the dashed red curve versus the wide range of dotted red curve). For simplicity, the variance of the spatial correlations at a given centroid distance was assumed to be the same for both subpopulations and was estimated based on the standard deviation of all neighboring cell-pairs and the estimated average spatial correlations of each subpopulation. To obtain the joint model, for each mouse individually, we modeled for any given centroid distance, the distribution of spatial correlations for the same cells as a lognormal distribution and for different cells as a beta distribution.

To obtain the probability for any pair of neighboring cells from different sessions, given their spatial correlation and centroid distance, we calculated  $P_{\text{same}}(\text{dist}, \text{corr})$ , according to Bayes' rule:

$$P_{\text{same}}(\text{dist}, \text{corr}) = \frac{P(\text{dist}, \text{corr} | \text{same}) W_{\text{same}}}{p(\text{dist}, \text{corr} | \text{same}) W_{\text{same}} + p(\text{dist}, \text{corr} | \text{diff}) W_{\text{diff}}}$$

where  $W_{\text{same}}$  and  $W_{\text{diff}}$  are the estimated weights (prior probabilities) for the two subpopulations of same cells and different cells, respectively.  $P(\text{dist}, \text{corr} | \text{same})$  and  $P(\text{dist}, \text{corr} | \text{diff})$  are the estimated conditional joint distributions of centroid distances and spatial correlations for same cells and different cells, respectively. Similar calculations were performed for each of the one-dimensional models (spatial correlation and centroid distance). This allowed us to obtain  $P_{\text{same}}$  for any pair of cells across sessions in the data.

### (3) Cell registration with a clustering procedure

Following the observation that cells are spatially clustered into clusters representing the same cells across imaging sessions, we applied an iterative clustering procedure aimed at maximizing within-cluster  $P_{\text{same}}$  and minimizing between-cluster  $P_{\text{same}}$ . To obtain a clustering of cells to serve as initial conditions for the iterative clustering procedure, we registered cells according to the spatial correlations model with a registration threshold of  $P_{\text{same}}=0.5$ . We created an initial list of cells comprising of the cells detected in the first session. Each session was then inspected and each cell that had a pairing candidate on the list, with which it had a  $P_{\text{same}}>0.5$ , was registered to that cell. Cells that did not have such a pairing candidate were added to the list as new cells. In cases where more than one candidate crossed the threshold, the cell-pair with the highest  $P_{\text{same}}$  was registered to be the same cell. The iterative clustering procedure was then applied. In each of the iterations, all cells were inspected and each cell either: (1) remained in the same cluster; (2) was transferred to a more similar cluster; or (3) initiated a new cluster. The decision was made by finding for each cell the candidate from all the different sessions with which it had the highest  $P_{\text{same}}$ . If the  $P_{\text{same}}$  was lower than the registration threshold, the cell formed a new cluster. If  $P_{\text{same}}$  was higher than the registration threshold, the cell joined the cluster of that candidate (if they were already in the same cluster, then nothing changed). If there was already a cell from that session with a lower  $P_{\text{same}}$  in the cluster, the new cell replaced the old one in the cluster, but if the existing cell had a higher  $P_{\text{same}}$ , then nothing changed. The iterative process continued until the clustering procedure converged and no more clustering changes occurred.

## Quantification of registration accuracy

### *Cell-pairs*

By applying  $P_{\text{same}}$  to our data, we obtained the certainty of registration associated with any pair of neighboring cells across sessions. We defined uncertain registrations of cell-pairs as

those with a probability  $\leq 0.95$  to be the same cell and  $\leq 0.95$  to be different cells ( $0.05 \leq P_{\text{same}} \leq 0.95$ ), and calculated the fraction of uncertain cell-pairs out of all neighboring cell-pairs. Additionally, we used the estimated probability distributions of same cells and different cells to calculate a receiver operating characteristic (ROC) curve obtained from 1,000 evenly spaced registration thresholds of  $P_{\text{same}}$  between 0 and 1. For the centroid distances model, the false negative rate was estimated as the fraction of the model of same cells that is higher than the registration threshold:

$$\text{False negative rate} = \frac{\# \text{false negatives}}{\# \text{false negatives} + \# \text{true positives}}$$

The false positive rate was estimated as the fraction of the model of different cells that is lower than the registration threshold:

$$\text{False positive rate} = \frac{\# \text{false positives}}{\# \text{false positives} + \# \text{true negatives}}$$

Similar calculations were performed for the spatial correlations and joint models. We defined the overall estimated registration accuracy:

$$\text{accuracy} = \frac{\# \text{true negatives} + \# \text{true positives}}{\# \text{false negatives} + \# \text{true positives} + \# \text{false positives} + \# \text{true negatives}}$$

Additionally, we calculated the Gini coefficient  $G_1$  to quantify the separability between the distributions of two subpopulations as  $G_1 = 2\text{AUC} - 1$ , where AUC is the area under the ROC curve.

### *Cell registers*

To evaluate the performance of our registration method for multiple sessions, we defined a stringent register score index for each registered cell, which takes into account false positives, false negatives, and non-exclusive cell registrations from all the pairwise combinations of registered sessions across all sessions. For each session, a cell was either active or inactive according to the cell registration, resulting in pairs of sessions where the cell was active in both (active-active), active in only one (active-inactive), or active in none (inactive-inactive — no evaluation is required in these cases). For each cell register in the data, a true positive score was calculated as the total number of reliable active-active cell-pairs ( $P_{\text{same}} > 0.95$ ) out of the total number of active-active pairwise combinations of all sessions. A true negative score was calculated as the total number of reliable active-inactive pairs ( $P_{\text{same}} < 0.05$  for all candidates) out of the total number of active-inactive pairwise combinations of all sessions. An exclusivity score was calculated as the total number of exclusive active-active cell-pairs (for all additional candidates  $P_{\text{same}} < 0.05$ ) out of the total

number of active-active pairwise combinations of all sessions. The register score was calculated as the total number of reliable pairs out of the total number of active-active and active-inactive pairwise combinations of all sessions, where each pair was considered reliable only if it met all three aforementioned criteria. Accordingly, we defined the register score as:

$$\text{register score} = \frac{1}{N(N-1)} \sum_{k=1}^N \sum_{\substack{m=1 \\ m \neq k}}^N \delta(k, m)$$

where N is the total number of sessions, N is the number of active sessions for a given cell register, and  $\delta(k, m)$  is 1 if a cell-pair in the cluster is reliable, and 0 otherwise.

#### *Scalability for a large number of sessions*

To examine if registration accuracy changes with the number of registered sessions, we applied our registration method to data recorded in 16 sessions on eight different days spanning two weeks (N=5 mice). Registration was performed simultaneously for multiple sessions with the clustering procedure, choosing random subsets of the 16 sessions comprising of 4, 8, or 12 sessions, and for the entire set of 16 sessions. The average register score was measured as a function of the number of registered sessions.

### **Validation of cell registration**

#### *Exclusivity and transitivity measures*

To validate our probabilistic approach, we defined *exclusivity* and *transitivity* measures. Exclusivity requires that only one cell from session 1 be paired with a given cell from session 2. Transitivity requires that if a cell from session 1 and a cell from session 2 are paired and the cell from session 2 is paired with a cell from session 3, then the cells from sessions 1 and 3 be also paired. To measure exclusivity, we calculated the distribution of  $P_{\text{same}}$  with all additional candidates only for cells that have a pairing candidate with a  $P_{\text{same}} > 0.5$  in another session. Cells with additional pairing candidates with a  $P_{\text{same}} > 0.5$  were considered non-exclusive. To measure transitivity, we calculated the distribution of  $P_{\text{same}}$  for pairs of cells from two different sessions where each had a  $P_{\text{same}} > 0.5$  with the same candidate from a third session. Of such cell-pairs, those with a  $P_{\text{same}} < 0.5$  were considered non-transitive. To compare exclusivity and transitivity of the data with those expected by chance, we obtained shuffled data by measuring the centroid distances between cell-pairs from different sessions, where each session was taken from a different mouse. We then performed the same analyses on the shuffled data.

### *Place field stability*

We also validated our probabilistic approach by applying our registration method to data recorded from the CA1 of the hippocampus of N=5 mice while freely exploring linear tracks (Rubin et al., 2015). We analyzed mouse behavior videos using a custom MATLAB (Mathworks) routine that detected the mouse's center of mass in each frame, calculated its velocity and applied a rectangular smoothing window of 250msec. For place field analysis, we considered periods wherein the mouse ran >3cm/sec. We divided each track into 24 bins (4cm each) and computed the time spent in each bin and the number of Ca<sup>2+</sup> events per bin, and smoothed these two maps ('occupancy' and 'Ca<sup>2+</sup> event number') using a truncated Gaussian kernel ( $\sigma=1.5$  bins, size=5 bins) (Ziv et al., 2013; Rubin et al., 2015). We then computed the activity map (event rate per bin) for each neuron by dividing the smoothed map of Ca<sup>2+</sup> event numbers by the smoothed map of occupancy. We separately considered place fields for the two running directions on the linear track. We defined each of the place field's position to be the peak value of the activity map. For each place field with  $\geq 5$  events for a given session, we computed the spatial information (in bits per event) using the unsmoothed events-rate map of each cell, as previously described (Markus et al., 1994):

$$\text{Spatial information} = \sum_i p_i (r_i / \bar{r}) \log_2 (r_i / \bar{r})$$

where  $p_i$  is the probability of the mouse to be in the  $i^{\text{th}}$  bin (time spent in  $i^{\text{th}}$  bin/total session time);  $r_i$  is the Ca<sup>2+</sup> event rate in the  $i^{\text{th}}$  bin;  $\bar{r}$  is the overall mean Ca<sup>2+</sup> event rate; and  $i$  is running over all the bins. We then performed 1,000 distinct shuffles of animal locations during Ca<sup>2+</sup> events, accounting for the spatial coverage statistics for the relevant session and direction, and calculated the spatial information for each shuffle. This yielded the p value of the measured information relative to the shuffles. Cells with spatial information higher than that of 95% of their shuffles were considered significant place cells. We then calculated the place field positional shift and place field correlation for cell-pairs from different sessions for which at least one cell in the pair was considered significant. The place field positional shift was calculated as the difference in place field peak positions, and the place field correlation was calculated as the Pearson correlation between the place field activity maps. The place field maps used for these two measures included both running directions on the running track. To compare the place field stability of the data with the stability expected by chance, the identity of cells were randomized (shuffled data) across sessions. We then performed the same analyses on the shuffled data.

### *Simulated data*

Simulated data was obtained by using representative spatial footprints of cells that were detected in our data. We then placed each representative spatial footprint in a random location in the FOV to serve as the original centroid location of the cell. We added a constraint that did not allow for two different cells to have very close centroid locations. The minimal centroid distance between the original locations of neighboring cells (different cells) was set to  $7\pm1\mu\text{m}$  (normally distributed) to match the distribution of the observed data. In each session, a fraction of the cells were simulated as active and their locations were set as the original centroid locations with a jitter. The radius of the jitter for each cell in each session was drawn from a lognormal distribution and the angle was uniformly distributed. The parameters for the active fraction of cells and for the jitter were chosen to match those of the observed data. To obtain simulated data with different noise levels, we systematically increased the radius of the jitter by increasing the mean of the lognormal distribution. We then applied our registration procedure to the simulated data. This allowed us to estimate the false positive and false negative rates from the simulated data model and compare them with the actual false positive and false negative rates obtained when applying either our cell registration method or a centroid distance threshold to the simulated data. Finally, we simulated the effects of registration errors on measurements of the stability of the functional properties of cells. We defined different cells (shuffled data) as having a certain probability ( $10\pm1\%$ ) to share the same coding properties across sessions. To test whether registration errors could obscure conclusions drawn from data with different population effect sizes, we systematically changed the percentages of same cells that maintain their coding properties across sessions ( $10-15\pm1\%$ ). Effect size was defined as the difference between the mean stability of the population and shuffled data, divided by the stability standard deviation.

## Supplemental Information References

- Bansal, N., Blum, A. and Chawla, S. (2004) Correlation Clustering, *Mach. Learn.*, 56, pp. 89–113.
- Chen, T. W. *et al.* (2013) Ultrasensitive fluorescent proteins for imaging neuronal activity, *Nature*, 499, pp. 295–300.
- Dana, H. *et al.* (2014) Thy1-GCaMP6 Transgenic Mice for Neuronal Population Imaging In Vivo, *PLoS One*, 9, p. e108697.
- Markus, E. J. *et al.* (1994) Spatial information content and reliability of hippocampal CA1 neurons: Effects of visual input, *Hippocampus*, 4, pp. 410–421.
- Mukamel, E. A., Nimmerjahn, A. and Schnitzer, M. J. (2009) Automated Analysis of Cellular Signals from Large-Scale Calcium Imaging Data, *Neuron*, 63, pp. 747–760.
- Pnevmatikakis, E. A. *et al.* (2016) Simultaneous Denoising, Deconvolution, and Demixing of Calcium Imaging Data, *Neuron*, 89, p. 299.
- Rubin, A. *et al.* (2015) Hippocampal ensemble dynamics timestamp events in long-term memory, *Elife*, 4, p. e12247.
- Zhou, P. *et al.* (2016) Efficient and accurate extraction of in vivo calcium signals from microendoscopic video data, *arXiv*, p. e1605.07266.
- Ziv, Y. *et al.* (2013) Long-term dynamics of CA1 hippocampal place codes., *Nat. Neurosci.*, 16, pp. 264–6.