Kilohertz-rate two-photon voltage imaging of population dynamics *in vivo*

Min Zhang[1,2#], Shoupei Liu[1,2#], Yaoguang Zhao[1,2], Yanfeng Zhu[1,2], Xinyang Gu[3,4], Cihang Kong[1], Jiahao Hu[1,2], Huiyun Yu[5], Jingchuan Wu[3,4], Fang Xu[6], Liang Chen[3,4], Ying Mao[3,4], Bo Li[2,7*]

[1]Jinshan Hospital, MOE Frontiers Center for Brain Science, Institutes for Translational Brain Research, Fudan University, Shanghai 200032, China

[2]School of Psychology, Institute of Chinese Cultural Psychology, Shanghai Jiao Tong University, Shanghai 200030, China.

[3]Department of Neurosurgery, Huashan Hospital, Shanghai Medical College, Fudan University, Shanghai 200040, China.

[4]National Center for Neurological Disorders, Shanghai, China.

[5]Department of Pharmacology, School of Basic Medical Sciences, Fudan University, Shanghai 200032, China.

[6]Interdisciplinary Center for Brain Information, Brain Cognition and Brain Disease Institute, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, 518055, China.

[7]Jinshan Hospital, Huashan Hospital, Institutes for Translational Brain Research, Fudan University, Shanghai 200032, China.

[#]These authors contributed equally: Min Zhang, Shoupei Liu.

*Correspondence: bo-li@sjtu.edu.cn (B.L.)

Abstract

Understanding how neural circuits compute requires capturing voltage dynamics across large neuronal populations with millisecond resolution *in vivo*. However, two-photon voltage imaging remains fundamentally limited by a trade-off among imaging speed, field of view, and excitation efficiency. We introduce HS2PM, a hybrid scanning two-photon microscope that overcomes this bottleneck, enabling kilohertz-rate imaging over a 650 × 524 μm^2 field of view at single-cell resolution while preserving the photon efficiency of single-point excitation. HS2PM stably records deep-layer membrane potentials from over 160 neurons simultaneously, with minimal photobleaching and phototoxicity. It resolves both spikes and subthreshold voltage dynamics *in vivo*, revealing how these jointly shape sensory adaptation and population coding. Beyond voltage imaging, HS2PM supports high-speed fluorescence lifetime and vascular flow imaging, establishing a multimodal platform for dissecting fast circuit dynamics with precision previously inaccessible to optical methods.

Introduction

Neural circuits compute at millisecond timescales, coordinating rapid voltage changes across distributed neuronal populations. Capturing these fast dynamics *in vivo* is essential for understanding how brain activity underlies behavior[1,2]. While calcium imaging has revolutionized population-scale neural recording[3,4], its inherently slow kinetics obscure the fast electrical signals that drive synaptic integration and action potentials[2,5,6].

Two-photon voltage imaging (TPVI) with genetically encoded voltage indicators (GEVIs) offers a path toward high-speed, deep-tissue recordings of membrane potential[7–12]. However, efforts to scale TPVI to large populations remain constrained by a three-way trade-off between imaging speed, field of view (FOV), and excitation efficiency. Parallelized approaches such as multi-point[7,13], line[14], or light-sheet[15,16] scanning increase acquisition rates by distributing excitation across space, but they sacrifice photon efficiency, increase laser power demand[17], and introduce cross-talk in scattering tissue[7]. In contrast, single-point scanning maximizes signal-to-noise ratio (SNR) and the number of sampled neurons[17–19], but it is limited by the speed of mechanical deflection systems[20–22].

To overcome these limitations, we developed HS2PM, a hybrid scanning two-photon microscope that integrates high-speed scanning with photon-efficient single-point excitation. This system achieves kilohertz-rate imaging across a 650 × 524 μm^2 FOV at single-cell resolution. HS2PM enables stable, high-SNR recordings from over 160 neurons in awake mice, reaching depths of 700 μm with minimal phototoxicity. It resolves both spikes and subthreshold voltage dynamics *in vivo*, revealing how they jointly contribute to sensory adaptation and population coding. In addition to voltage imaging, HS2PM supports high-speed fluorescence lifetime and

vascular flow imaging, offering a versatile platform for multimodal interrogation of circuit dynamics at millisecond resolution.

Results

Overcoming the speed-efficiency bottleneck

TPVI is fundamentally limited by a trade-off between excitation efficiency and imaging speed. Single-point scanning provides high efficiency but limited speed, whereas parallel strategies improve speed at the expense of per-focus efficiency and fidelity in scattering tissue. This efficiency loss is particularly detrimental for voltage imaging, where the photon budget is already tight due to membrane-localized GEVIs and millisecond-scale dynamics. A theoretical analysis (Supplementary Note 1) showed that, even under ideal single-point conditions, voltage imaging would require more than 50-fold greater laser power than calcium imaging to achieve comparable discriminability index (d'). When parallel strategies further divide excitation across multiple foci, the effective power requirement multiplies[17], rapidly driving the system beyond safe limits ($\leqslant$250 mW at 920 nm[23]). Together, these considerations establish that voltage imaging fundamentally requires single-point excitation to maximize efficiency, while speed gains must instead be realized through innovations in scanning architecture.

HS2PM overcomes the speed bottleneck of single-point scanning, achieving kilohertz-rate imaging across a large population-scale FOV. This is enabled by a hybrid scanning architecture that combines fast angular encoding with high line-rate deflection (Fig. 1a; Extended Data Fig. 1). First, an electro-optic deflector (EOD) angularly encoded the 80 MHz pulse train, cycling every four pulses to generate four spatiotemporally indexed points (Fig. 1b). These were further expanded by a spatiotemporal division multiplexing (STDM) module, which split each pulse into four time-staggered replicas, yielding 16 spatiotemporally indexed points at an effective sampling

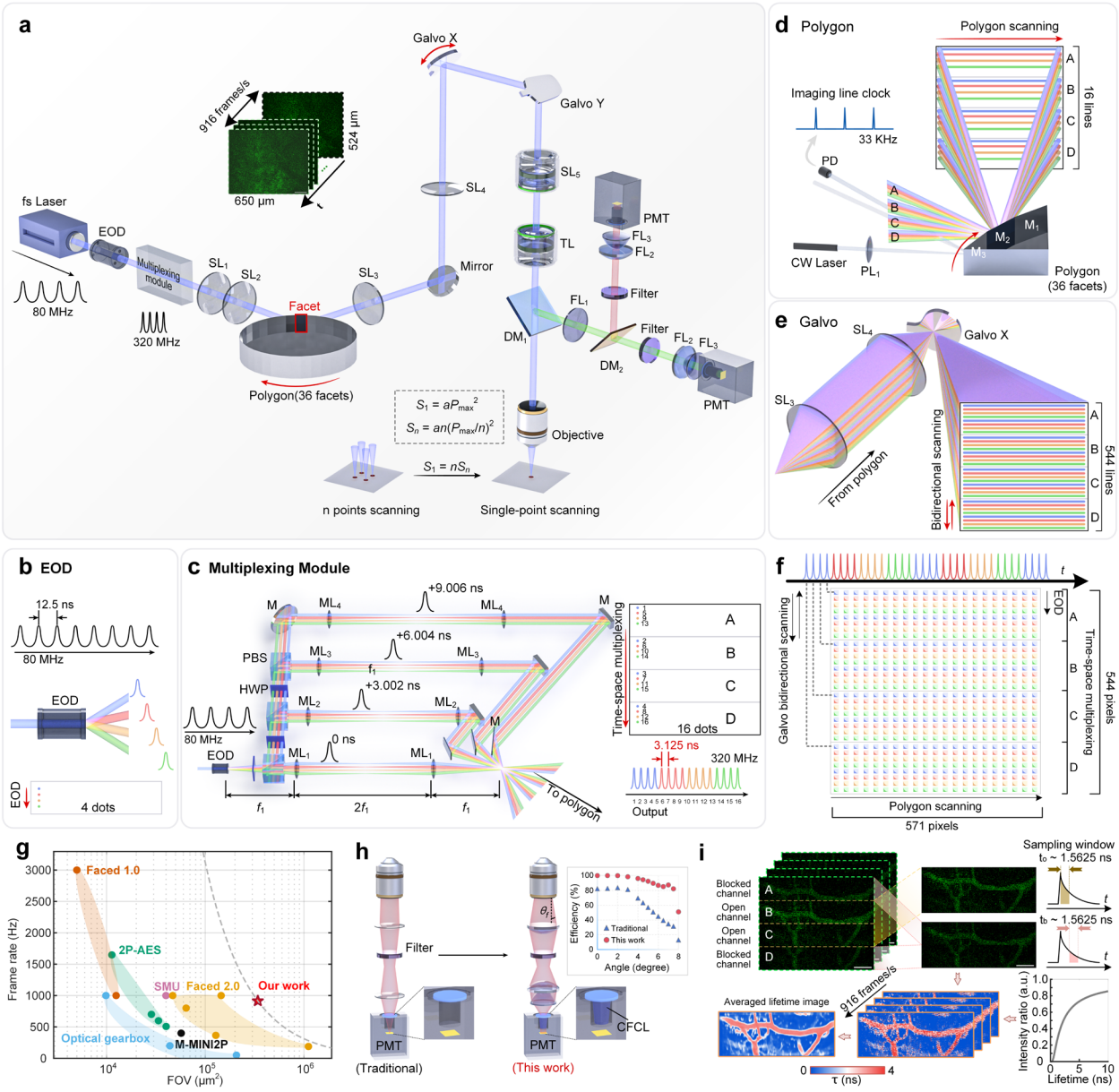


Fig. 1 Hybrid scanning architecture and optical design of HS2PM. a, Optical layout of the HS2PM system. An 80 MHz femtosecond (fs) laser (920 nm) beam is first angularly encoded by an EOD. It is then split into four time-delayed replicas by the STDM module and scanned along the *x*- and *y*-axes by a polygon and galvo scanner, respectively. This enables 916 Hz raster imaging over a 650 × 524 μm² FOV, with fluorescence collected via a dual-channel wide-angle detection system. **b**, Angular encoding by the EOD generates four spatially offset pulses from the 80 MHz pulse train. **c**, Polarization beam splitters (PBSs), half-wave plates (HWPs), and achromatic delay lines in the STDM module replicate each pulse into four delayed beams (~3 ns spacing). **d**, The polygon scanner scans all 16 beams to produce 16 lines per facet. A continuous-wave (CW) laser and photodiode (PD) generate synchronization clocks for scan alignment. **e**, Bidirectional galvo scanning fills the *y*-axis gaps between four-line clusters, producing 544 lines per frame.

f, Excitation pattern per frame, covering the full FOV. **g**, Comparison with other single-point scanning TPVI systems shows HS2PM achieves the largest FOV at kilohertz frame rates. Data from published studies (circles) with hypothetical interpolation trends (clouds). **h**, Comparison of emission collection strategies. The custom-designed collection system achieves higher collection efficiency than conventional epi-fluorescence relays within ± 8°. **i**, FLIM mode. STDM channels are reconfigured to generate two temporally offset intensity windows per pixel. Lifetime is estimated from their intensity ratio using a pre-calibrated lookup table.

clock of 320 MHz (Fig. 1c). A 36-facet polygon scanner rotating at 54,945 rpm then provided a base line rate of ~33 kHz. Scanning the 16 indexed points generated 16 lines, with an effective line rate of 527 kHz (Fig. 1d). Finally, a bidirectionally driven galvo performed *y*-axis scanning to generate 544 lines per frame (Fig. 1e). Together, this architecture enabled 916 Hz across 571 × 544 pixels (650 × 524 μm^2 FOV; Fig. 1f; see also Extended Data Fig. 2 for pixel order map). Among reported fast point-scanning systems capable of voltage imaging[12,19,24–27], HS2PM achieves a kHz frame rate over the largest FOV to date (Fig. 1g; Supplementary Table 1).

HS2PM substantially improves photon detection efficiency to boost SNR and depth penetration. Conventional epi-fluorescence relays, typically comprising one or two lenses[7,12,28], capture only a small fraction of the multiply scattered emission emerging from the objective back aperture at large angles. In HS2PM, we replaced this relay with a three-lens condenser coupled to a custom frustum-of-cone light guide (CFCL), which improves the photon collection efficiency within ± 8° and achieves an overall detection efficiency >80% (Fig. 1h; Extended Data Fig. 3). This configuration substantially increases the number of scattered photons reaching the detector, thereby improving SNR and extending the usable imaging depth at a given excitation power.

HS2PM can also be reconfigured for high-speed fluorescence lifetime imaging (FLIM), enabling *in vivo* estimation of indicator fluorescence lifetimes. In this mode, the STDM output is configured to two active channels, and for each pixel we acquire two temporally separated

115 sampling windows: one corresponding to the prompt emission and the other delayed by
116 approximately 1.56 ns (Fig. 1i; Extended Data Fig. 4b). The fluorescence lifetime at each pixel is
117 then derived from the ratio of the two intensity values, based on the exponential decay kinetics of
118 the fluorophore. Although disabling two STDM channels reduces spatial throughput and halves
119 the accessible FOV, this FLIM mode enables quantitative *in vivo* estimation of indicator
120 fluorescence lifetimes, broadening HS2PM's functionality beyond intensity imaging.

121 HS2PM integrates a streamlined data-processing pipeline that preserves millisecond
122 dynamics and enhances neuron identifiability at scale (Extended Data Fig. 5). Raw frames are first
123 rearranged to restore raster images using a pixel-wise order map (Extended Data Fig. 2), with an
124 additional correction for odd-even offsets from bidirectional galvo scanning. STDM crosstalk is
125 then demixed to suppress signal contamination when fluorescence lifetimes exceed the sampling
126 interval (see Methods). Finally, VolPy[29] is used for motion correction, denoising and robust spike
127 extraction. This end-to-end pipeline ensures high-fidelity voltage readout, enabling reliable
128 population-scale imaging with millisecond precision.

129 **Stable and deep-layer population voltage imaging with minimal phototoxicity**

130 HS2PM stably resolves millisecond-scale voltage dynamics from large populations of cortical
131 neurons *in vivo*, overcoming the long-standing difficulty of maintaining high-SNR recordings
132 hundreds of micrometers below the cortical surface[6]. Using a soma-targeted voltage indicator
133 (ASAP5-Kv) in awake, head-fixed mice, we imaged a 650 × 524 μm^2 FOV with a 16× 0.8 NA
134 objective, identifying 165 neurons in a single plane at 490 μm depth under 200 mW excitation
135 (Fig. 2a). Voltage traces from these neurons exhibited robust action potentials with peak $-\Delta F/F$
136 amplitudes of ~35% on average (Fig. 2b). Low-pass filtering revealed subthreshold fluctuations,

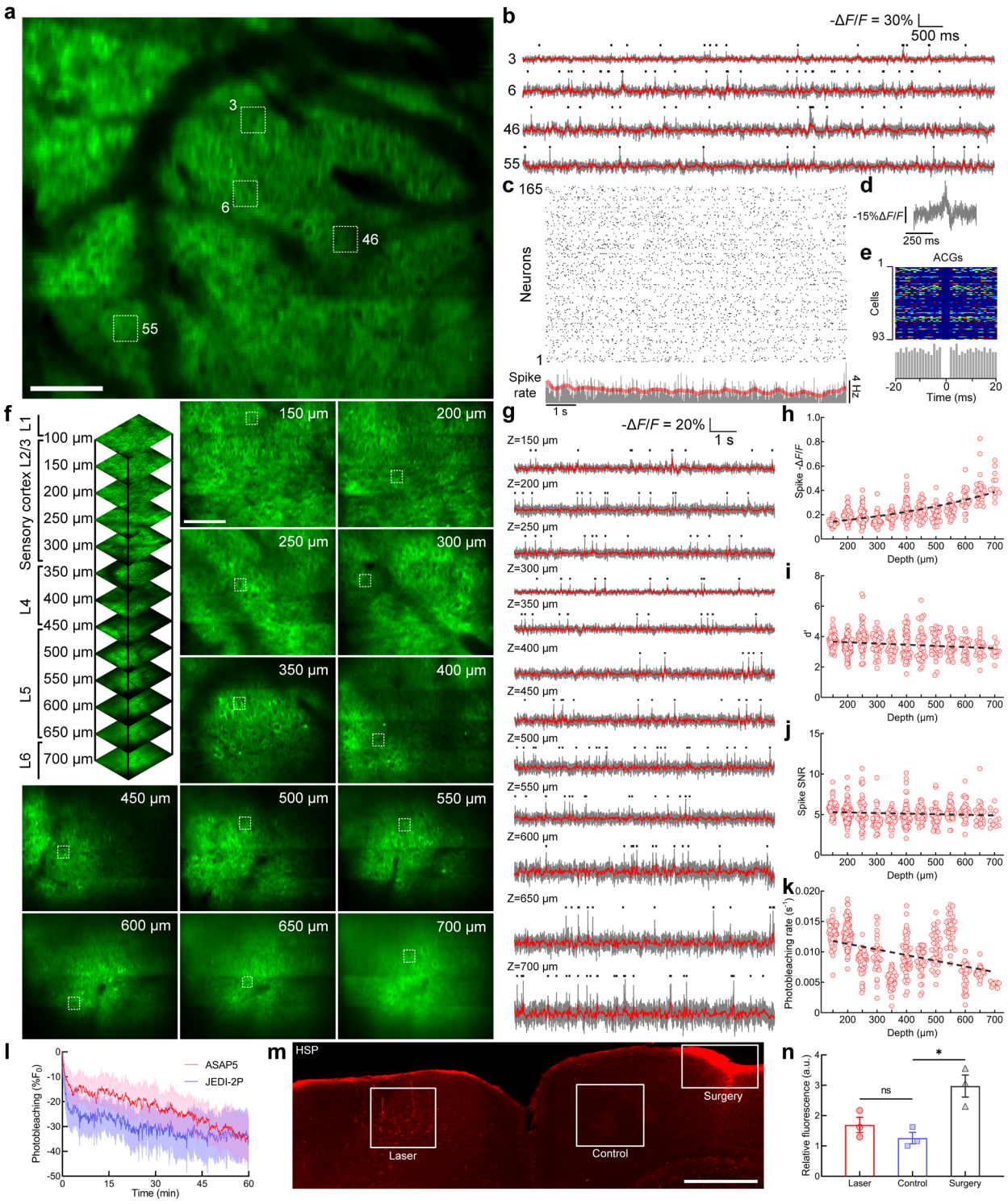


Fig. 2 HS2PM enables stable, deep-layer, and long-duration population voltage imaging in awake mouse cortex. a, Time-averaged projection of an *in vivo* ASAP5-Kv-labeled image recorded at 490 µm depth in sensory cortex. FOV: 650 × 524 µm^2 (571 × 544 pixels), acquired at 916 Hz using 920 nm excitation at 200 mW post-objective (Nikon 16×, 0.8 NA). Scale bar, 100 µm. **b**, Representative -Δ*F*/*F*

142 voltage traces (gray) from indexed neurons in **a**, showing robust spike events (black ticks) and subthreshold
143 fluctuations (red). **c**, Raster plot of detected spikes from 165 simultaneously imaged neurons, aligned with
144 the average population firing rate (gray) and its low-pass filtered trace (red). **d**, Example of an individual
145 spike waveform captured in raw fluorescence, exhibiting high temporal resolution and clear peak structure.
146 **e**, ACGs of neurons with ≥20 spikes (each row), indicating the ± 2.2 ms ISI. Bottom histogram: population-
147 averaged ACG. **f**, Time-averaged projection of ASAP5-Kv-labeled neurons across cortical depths (150-700
148 μm) captured in a 373 × 300 μm^2 FOV using an Olympus 25×, 1.05 NA objective. Image size: 571 × 544
149 pixels. Scale bar, 100 μm. **g**, Representative -$\Delta F/F$ voltage traces (gray) and subthreshold components (red)
150 from neurons at various depths in **f**. **h**-**k**, Quantification of spike -$\Delta F/F$, d', spike SNR and bleaching rate.
151 **l**, One-hour continuous imaging using ASAP5 (n = 61 neurons, 2 FOVs, 2 animals) and JEDI-2P (n = 55
152 neurons, 2 FOVs, 2 animals) shows stable performance with < 35% bleaching under 200 mW excitation.
153 Shaded region, s.d. **m**, Immunostaining of HSP-70/72 to assess photodamage. Three cortical regions (white
154 box)—the imaged area, contralateral control region, and surgical injury site—show no significant damage
155 in the imaged area. Scale bar, 500 μm. **n**, Quantification of relative HSP-70/72 expression in imaged and
156 injured regions, normalized to contralateral control (n = 3 animals; Welch's t test; $*p < 0.05$, ns = not
157 significant). Error bars: s.e.m.

158 including slow-wave dynamics distinct from suprathreshold spikes. At the population level, spike
159 raster plots revealed coordinated activity (Fig. 2c, top), and the population-averaged firing rate
160 exhibited rhythmic fluctuations (Fig. 2c, bottom). We analyzed population activities from one of
161 our experimental animals, which exhibited more distinct synchronized spiking that was tightly
162 phase-locked to the oscillation cycles (Extended Data Fig. 6). Similar membrane oscillations have
163 previously been observed from neurons in sensory cortex[30]. Single-neuron traces exhibited
164 stereotyped action potential waveforms, with a rapid depolarization followed by a clear after-
165 hyperpolarization, consistent with refractory periods (Fig. 2d). Autocorrelogram (ACG) analysis
166 across the recorded neurons yielded inter-spike interval distributions of approximately ± 2.2 ms
167 (Fig. 2e), confirming the millisecond precision of HS2PM recordings *in vivo*.

168 HS2PM also maintains high-fidelity voltage recordings across cortical layers, achieving
169 reliable spike detection down to 700 μm depth. With a higher NA objective (25×, 1.05 NA; FOV

170 373 × 300 μm^2), we captured population-level voltage dynamics spanning layers 1-5 and part of
171 layer 6 (Fig. 2f). Voltage traces from neurons at all depths exhibited robust action potentials (Fig.
172 2g). Spike -Δ*F*/*F* amplitudes remained consistently high across depths (≥22% on average; Fig. 2h)
173 and showed a slight increase with depth. Spike detection was highly reliable across depths.
174 Although *d′* declined modestly and quasi-linearly from 150 μm to 700 μm, the mean value
175 remained ~3.5, and at 700 μm it was still >3 (Fig. 2i), corresponding to a high probability of
176 detection[31]. Spike SNR was largely preserved (~5) across cortical layers (Fig. 2j), and bleaching
177 rates remained low <0.012 s^{-1} (Fig. 2k). These results confirm that HS2PM preserves reliable spike
178 detection and signal fidelity in deep cortical layers.

179 Finally, HS2PM enables long-duration population imaging with minimal bleaching and
180 negligible phototoxicity, overcoming a critical barrier to extended *in vivo* voltage recordings.
181 Continuous 1-h recordings at 200 mW in six mice (three expressing ASAP5-Kv and three JEDI-
182 2P) produced only a modest ~35% fluorescence decline (Fig. 2l), indicating stable performance
183 with limited bleaching. To further assess phototoxicity, we examined HSP-70/72 expression.
184 Immunostaining revealed no detectable increase in the imaged regions compared to contralateral
185 controls (n = 3, t = 1.385, p = 0.2440, Welch's t test; Fig. 2m, 2n), whereas strong staining was
186 evident at surgically injured sites caused by head restraint (n = 3, t = 4.194, p = 0.0248). Together,
187 these results establish HS2PM as a robust platform for high-fidelity, long-duration voltage imaging
188 *in vivo*, combining stable population recordings with minimal bleaching and negligible
189 phototoxicity.

190 **Resolving spikes and subthreshold dynamics during sensory processing**

191 Large-scale calcium imaging has revealed transient reorganizations of cortical population activity
192 during sensory responses[32]. Yet the slow kinetics of calcium indicators obscure the millisecond

193 voltage transients that drive these dynamics, leaving the relative roles of spikes and subthreshold
194 fluctuations unresolved. Using HS2PM, we directly imaged *in vivo* voltage responses to sensory
195 stimulation, capturing both suprathreshold spiking and subthreshold membrane dynamics with
196 millisecond precision.

197 HS2PM revealed that sustained sensory stimulation drives a rapid and coordinated spiking
198 response across cortical populations. We applied a 3-s continuous air puff (Fig. 3a, top) to head-
199 fixed mice and imaged neuronal voltage dynamics in sensory cortex with a 25× 1.05 NA objective
200 at 355 μm depth (Fig. 3b; 200 mW). Analysis of 9-s voltage traces showed that responsive neurons
201 generated nearly synchronous spikes at stimulus onset (Fig. 3c). This activation was reflected in
202 normalized spike-rate heat maps from 297 neurons in two mice (Fig. 3d), where population-
203 averaged firing rates rose abruptly to ~9.5 Hz and peaked ~62.6 ms post-stimulus before relaxing
204 to baseline. Response-latency mapping (Fig. 3e) revealed substantial cell-to-cell variability, yet all
205 activations occurred within 223 ms. Across the entire population, latency histograms exhibited a
206 unimodal peak at 115.6 ms (Fig. 3f), consistent with rapid neuronal adaptation and feed-forward
207 drive[33]. Statistical analyses confirmed that, compared to the period without air puff, the air-puff
208 period showed both an overall increase in spike rate (Fig. 3g) and a modest rise in pairwise
209 correlations (Fig. 3h). Finally, three-dimensional principal component analysis (PCA) trajectories
210 showed a clear transition from pre- to post-stimulus states (Fig. 3i), indicating a pronounced
211 reorganization of cortical activity.

212 Repeated stimulation revealed a dissociation between spiking adaptation and stable
213 subthreshold dynamics. We applied a 10 × 1-s air puffs within 20 s (Fig. 3a, bottom), across five
214 trials per FOV. Each puff evoked an immediate spike-rate increase, with the strongest response

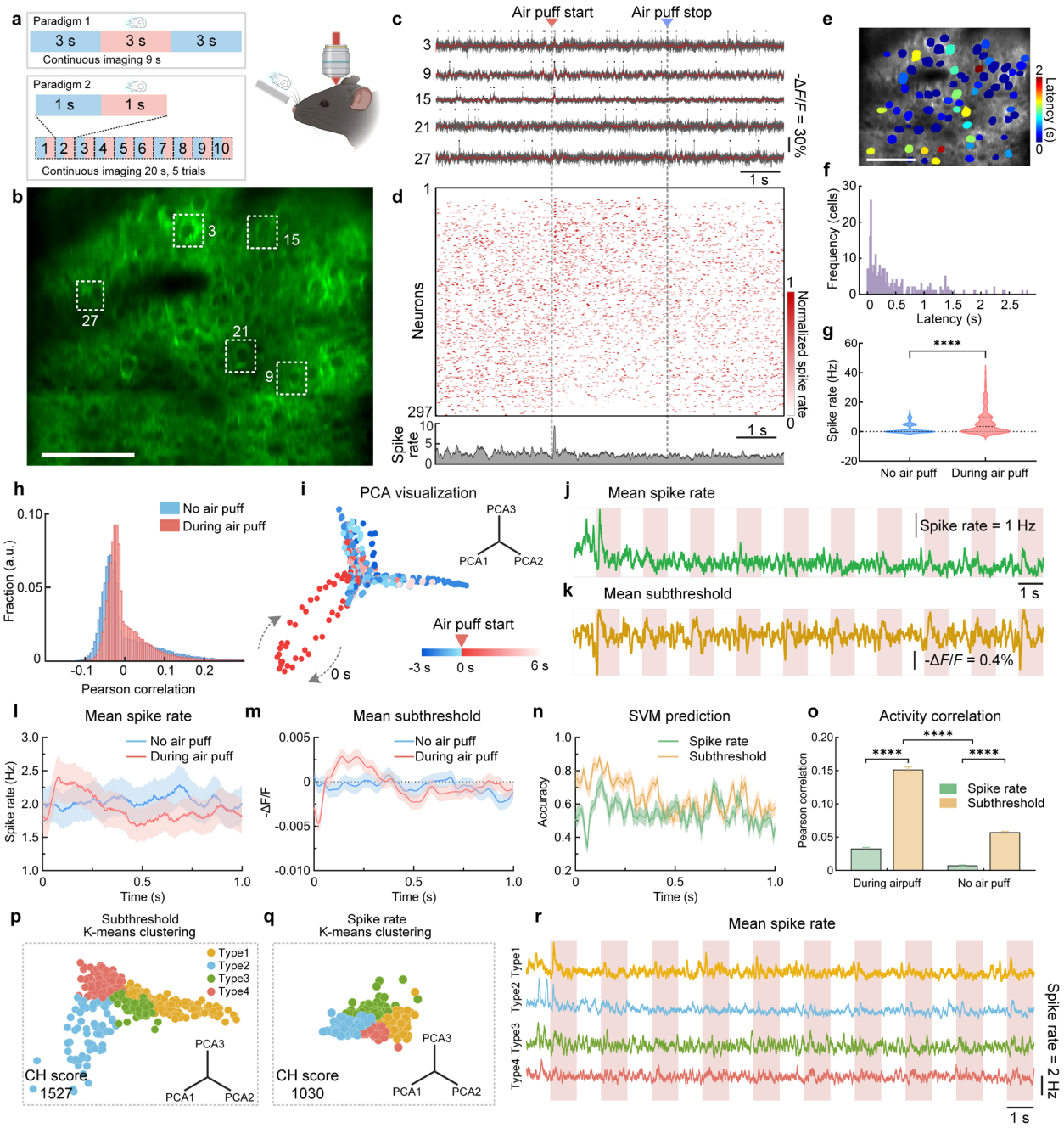


Fig. 3 HS2PM resolves rapid cortical voltage dynamics during sensory processing. a, Schematic of the two air-puff stimulation paradigms. **b**, Time-averaged image of neurons expressing ASAP5 in sensory cortex (355 µm depth) acquired at 916 Hz (25×, 1.05 NA). **c**, Representative voltage traces from indexed neurons in **b**, showing spikes (black ticks) and subthreshold signals (red) aligned to stimulus onset. **d**, Normalized spike rates of 297 neurons from 2 mice reveal rapid, synchronized responses following stimulus onset; population average shown below. **e**, Latency map of responsive neurons color-coded by time to first evoked spike; see also **f**, histogram of response latencies (peak at ~116 ms). **g**, Spike rates significantly

increased post-stimulation (****$p < 0.0001$, paired t-test). **h**, Pairwise Pearson's correlation coefficients of spike trains also increased after stimulation. **i**, PCA of population spiking activity revealed a clear trajectory shift following stimulus onset. **j-k**, Repeated stimuli evoked robust initial spiking (**j**) but stable subthreshold responses across trials (**k**). **l-m**, Both spikes and subthreshold signals increased during stimulation and returned to baseline. **n**, Classifiers trained on subthreshold signals outperformed those based on spike rates in decoding stimulus identity (SVM accuracy: 64% vs. 55%). **o**, Air puff increased neuron-population synchrony in both spike rate and subthreshold domains (****$p < 0.0001$, 2-way ANOVA). **p-q**, K-means clustering based on subthreshold dynamics yielded more cohesive and interpretable response types (Calinski-Harabasz score: 1527 vs. 1030). **r**, Subthreshold-based clusters exhibited distinct adaptation patterns across repeated stimuli: Type 1 neurons remained responsive, while Types 2/3 adapted and Type 4 was non-responsive. Scale bars: 100 μm in **b** and **e**. Shaded region in **l-n**, s.e.m.

during the first puff and progressively attenuated activity thereafter (Fig. 3j). In contrast, subthreshold voltage traces exhibited reproducible -$\Delta F/F$ transients for every puff, with amplitudes remaining stable across repetitions (Fig. 3k). At the population level, both spike rates and subthreshold signals rose physically at puff onset and decayed back to baseline (Fig. 3l and 3m), confirming consistent stimulus encoding across suprathreshold and subthreshold domains. Quantitative analyses further highlighted the informational advantage of subthreshold signals: support vector machine (SVM) classifiers trained on subthreshold traces outperformed those based on spike-rate vectors (64.01 ± 4.42 % vs. 54.61 ± 4.54 %, Fig. 3n), and both spike and subthreshold correlations with the surrounding population increased significantly during stimulation (r_{rate} = 0.0327 ± 0.0019 vs. 0.0074 ± 0.0008; r_{sub} = 0.1515 ± 0.0038 vs. 0.0573 ± 0.0016; $p < 0.0001$; Fig. 3o). Together, these results demonstrate that while spikes adapt under repeated stimulation, subthreshold dynamics remain robust, supporting enhanced network-level coordination.

Subthreshold activity provided a more reliable basis for classifying functional response types. We applied unsupervised k-means clustering to neurons grouped by either spike-rate profiles or subthreshold voltage dynamics. Both approaches partitioned the population into four functional

classes. Subthreshold-based clustering produced higher intra-cluster cohesion, as quantified by the Calinski-Harabasz (CH) score (1527 vs. 1030; Fig. 3p,q), and produced distributions that were more consistent with biologically interpretable response types (Type 1-4: n_1 = 210, n_2 = 82, n_3 = 60, n_4 = 233 in Fig. 3p vs. n_1 = 158, n_2 = 254, n_3 = 97, n_4 = 76 in Fig. 3q). We therefore adopted the subthreshold classification for subsequent analyses. The four classes exhibited distinct response patterns to repeated air puff stimuli (Fig. 3r): Type 1 neurons responded robustly to every puff, Type 4 neurons remained largely inactive, and Types 2 and 3 showed strong initial responses that rapidly attenuated with repetition. These results demonstrate that subthreshold-based clustering yielded higher fidelity and revealed functional heterogeneity in neuronal adaptation to repeated stimuli.

Together, HS2PM resolves millisecond-scale reorganization of cortical activity and reveals how spikes and subthreshold dynamics jointly contribute to population coding, providing insights into sensory processing that remain inaccessible with calcium imaging.

Population-scale vascular imaging and high-speed fluorescence lifetime imaging

HS2PM enables population-scale vascular imaging by overcoming the one-vessel limitation of conventional two-photon line-scanning approaches[34,35]. Whereas traditional methods restrict flow measurements to a single vessel, the near-kHz frame rates of HS2PM provide the temporal resolution needed for high-throughput analysis of flow velocity and diameter across many vessels simultaneously. To demonstrate this capability, we labeled the cerebral vasculature in mice with Rhodamine-B and recorded vascular fluorescence at 916 Hz, using a 25× objective at 90 μm depth (100 mW; Fig. 4a). Millisecond-resolved traces from two representative vessel segments revealed flow fluctuations with average velocities of 0.0875 mm/s and 0.0935 mm/s, respectively (Fig. 4b).

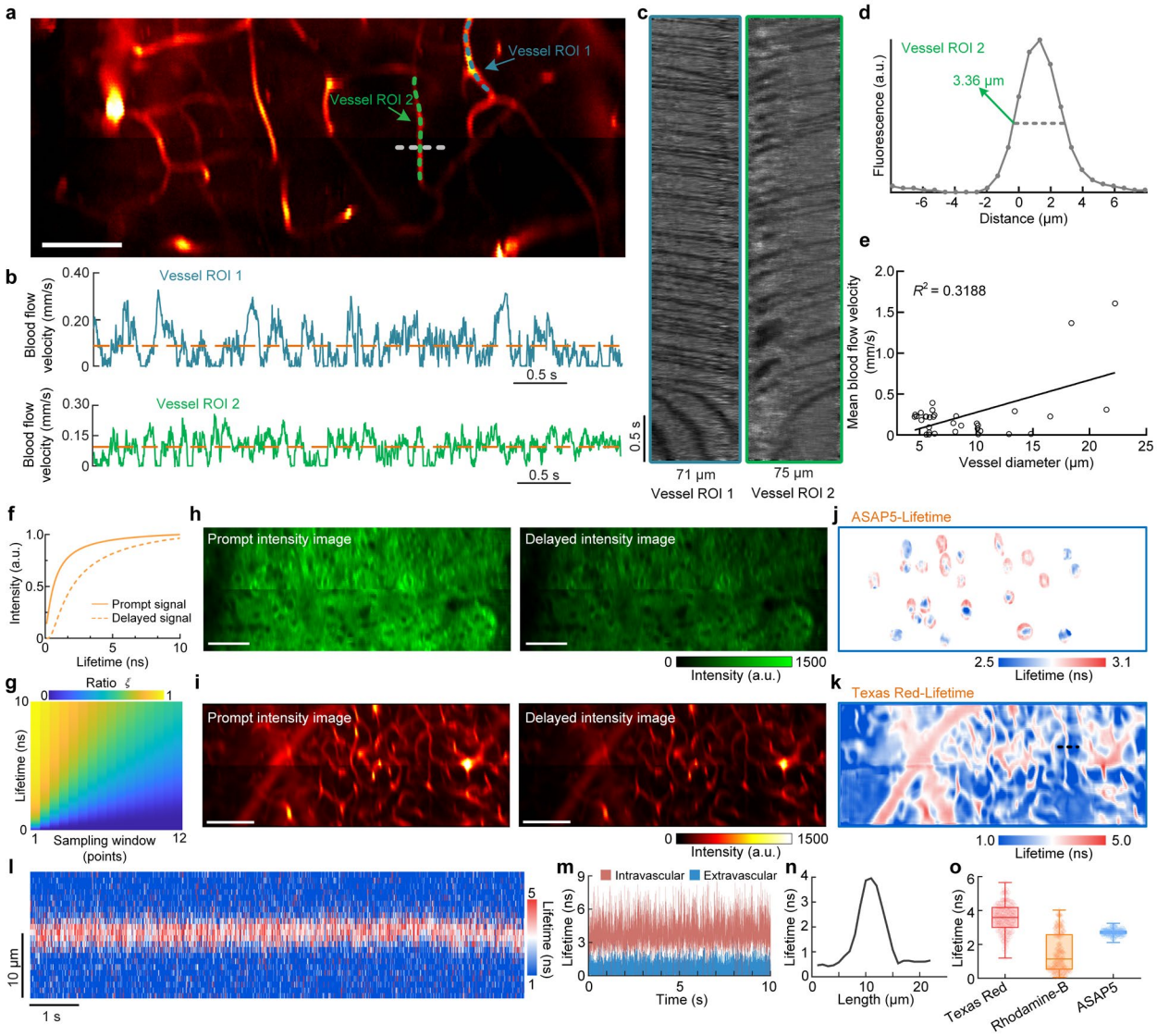


Fig. 4 HS2PM enables high-speed population-scale vascular imaging and pixel-resolved fluorescence lifetime mapping. a, Time-averaged projection of vasculature labeled with Rhodamine-B at 100 μm depth. Two vessel segments (ROIs) are traced for flow analysis. **b**, Flow velocity fluctuations over 5 s from the two vessel ROIs. **c**, Kymographs showing RBC movement as diagonal streaks. **d**, Vessel diameter extracted by Radon transform and FWHM. **e**, Positive correlation between vessel diameter and flow velocity (n = 33 vessels). **f-g**, FLIM calibration curves and color-coded map linking fluorescence lifetime to intensity ratio (ξ) across dual time-offset channels (see methods for details). **h-i**, Prompt and delayed emission frames of ASAP5-expressing neurons (185 μm depth) and Texas Red-labeled vasculature (90 μm depth). **j-k**, Corresponding pixel-resolved lifetime maps. **l-n**, Lifetime contrast across vessel cross-section reveals elevated lifetime within lumen; fluctuations and spatial profile are consistent with erythrocyte flux. **o**,

Summary of lifetime measurements across indicators: Texas Red, Rhodamine-B, and ASAP5 (n = 86, 86, 58 ROIs). Scale bars: 50 μm in **a** and **h**, 100 μm in **i**.

Spatiotemporal kymographs visualized red blood cell movement as diagonal streaks (Fig. 4c), while vessel diameter was extracted from fluorescence cross-sections using Radon transform and defined as the full-width at half-maximum (FWHM; Fig. 4d). Across 33 vessel segments, HS2PM enabled simultaneous quantification of flow velocity and vessel diameter, revealing a positive correlation between the two metrics (R^2 = 0.3188; Fig. 4e), consistent with physiological expectations. These results establish the feasibility of population-scale vascular flow measurements with HS2PM.

HS2PM extends beyond intensity imaging by enabling high-speed, pixel-resolved FLIM. We implemented a dual-channel ratio method in which fluorescence was acquired in two temporally offset sampling windows, corresponding to prompt and delayed emission (Extended Data Fig. 4b). The ratio ξ between these signals directly encodes fluorescence lifetime based on exponential decay kinetics (Fig. 4f), thereby converting conventional intensity images into quantitative, pixel-resolved lifetime maps (Fig. 4g). Applying this approach, we resolved lifetime contrasts in both neurons and vasculature. ASAP5-labeled neurons and Texas Red-labeled vessels produced distinct peak and faded intensity images (Figs. 4h,i), which were converted into pixel-resolved lifetime maps (Figs. 4j,k). In vessel cross-section, lifetime mapping revealed clear contrasts between intravascular and extravascular regions: the vessel lumen exhibited fluctuating lifetimes, whereas surrounding tissue showed consistently short values (~0.58 ns; Fig. 4l,m,n), consistent with endogenous autofluorescence. Finally, quantitative analysis across regions of interests (ROIs) confirmed robust fluorophore-specific lifetimes: Texas Red- and Rhodamine-B-labeled vasculature exhibited mean values of 3.57 ± 0.07 ns and 1.50 ± 0.10 ns, respectively, while

305 ASAP5-expressing neurons showed 2.72 ± 0.01 ns (mean ± s.e.m.; Fig. 4o), consistent with prior
306 benchmarks[36–38].

307 Discussion

308 We present HS2PM, a hybrid scanning two-photon microscope that resolves the long-standing
309 trade-off between imaging speed, FOV, and excitation efficiency in TPVI. By preserving the
310 photon efficiency inherent to single-point excitation while achieving kilohertz frame rates across
311 a population-scale FOV, HS2PM enables reliable *in vivo* recording of both spiking and
312 subthreshold voltage dynamics with single-cell and one-millisecond resolution.

313 This study introduces a conceptual and technical framework that redefines the design
314 constraints of TPVI. Through theoretical modeling, we demonstrate that the fundamental
315 limitation lies not only in scanning speed but also in photon flux under the biological power ceiling.
316 Our analysis shows that within realistic power limits, single-point excitation offers the most
317 practical route to achieving sufficient excitation efficiency for reliable voltage detection
318 (Supplementary Note 2). This underscores that excitation efficiency is as critical as frame rate.
319 Based on this insight, we developed a hybrid scanning strategy that overcomes the intrinsic speed
320 limit of point scanning. Operating near the theoretical upper bound imposed by fluorophore
321 lifetimes of approximately 3 nanoseconds, HS2PM achieves kilohertz-rate imaging at 320 MHz
322 effective sampling without compromising FOV. The resulting square-format population-scale
323 imaging field allows seamless adaptation of established calcium imaging paradigms to voltage
324 imaging. In addition, the same temporal multiplexing framework supports high-speed fluorescence
325 lifetime imaging, enabling kilohertz-rate lifetime mapping without requiring electro-optic gating[39]
326 while also integrating deep, fast, and quantitative voltage imaging within a single system.

By scaling voltage imaging to the population level, HS2PM enables circuit-level analyses that were previously limited to calcium imaging, but with millisecond precision sufficient to resolve both action potentials and subthreshold membrane fluctuations *in vivo*. This capability allows direct dissection of how spikes and subthreshold dynamics jointly shape sensory adaptation, network coordination, and population coding. These results provide new experimental access to fundamental questions in systems neuroscience, for example how neural circuits compute[40], how subthreshold activity contributes to collective dynamics[41,42], and how excitatory and inhibitory inputs interact to shape cortical responses[42], questions that have remained inaccessible with calcium imaging alone.

Several limitations remain. The fluorescence lifetime of existing GEVIs (~3 ns)[38] restricts further gains in temporal multiplexing and introduces signal crosstalk[43], necessitating demixing algorithms for correction. In addition, the 36-facet polygon scanner imposes a physical constraint on FOV, though recent two-pass geometries suggest this could be expanded up to twofold[27]. Biologically, variable GEVI expression levels and the absence of demonstrations in chronic or freely moving models remain critical challenges to broader application.

Future directions include pairing HS2PM with next-generation voltage indicators that offer shorter lifetimes, enhanced brightness, or red-shifted spectra. This would further improve imaging speed, depth, and multiplexing capabilities. Beyond neuronal populations, HS2PM may be applied to vascular dynamics, neuron-glia interactions, and all-optical physiology. The system provides a flexible platform for multimodal interrogation of neural circuits. With continued development, HS2PM has the potential to achieve population-scale millisecond-resolved imaging of brain activity in both healthy and diseased states. This positions voltage imaging to fulfill a

349 transformative role in neuroscience research comparable to that of two-photon calcium imaging
350 over the past two decades.

351 Methods

352 HS2PM microscope

353 We simulated and optimized the placement of optical components using Zemax OpticStudio
354 software (Zemax LLC). The complete design was then assembled in a 3D SolidWorks
355 environment to verify there was no component interference. Custom aluminium holders and
356 frames were machined for rigid mounting of the optics. A detailed schematic of the HS2PM setup
357 is shown in Extended Data Fig. 1 (see the animated rendering in Supplementary Video 1), with
358 components listed in Supplementary Table 2. A 920-nm, 2-W fiber laser (ALCOR920, Sparks
359 Lasers; 100-fs pulse duration, 80-MHz repetition rate) was used. The laser spot was focused onto
360 the sample via optical scanning devices, relay lenses and an objective (CFI75 LWD 16X W, Nikon;
361 16×/0.8 NA). The $1/e^2$ beam diameter at the back aperture is ~5.31 mm, leading to an excitation
362 NA of ~0.28. A 1.05 NA objective (XLPLN 25xWMP2, Olympus; 25×) was used for higher-
363 resolution voltage imaging with an excitation NA of ~0.46. The post-objective power was
364 controlled by the acousto-optic modulator (AOM) integrated in the laser.

365 The EOD (Model 400-120, Conoptics; 8.5 μrad/V deflection efficiency) was driven by an
366 amplifier (Model 25A, Conoptics) synchronized to the laser clock. An arbitrary waveform
367 generator (AWG; including PCI-8361, PXIe-8360 and PXIe-1062Q; National Instruments) output
368 20 MHz periodic waveforms, switching the voltage applied to the EOD every 12.5 ns. By cycling
369 through four distinct voltage levels, the EOD deflected successive laser pulses into four
370 longitudinal positions. These beams then entered the multiplexing module, which further split each
371 into four spatially separated channels via a combination of time- and space-division multiplexing.

This produced four vertical excitation points per scanning region, totaling 16 aligned points across the full FOV.

The multiplexing module consisted of three half-wave plates (HWP; 2 × WPH05M-915, 1 × WPH10M-915, Thorlab), three polarizing beam splitter cubes (PBS; PBS105, Thorlabs), eight lenses, and several mirrors. Three HWPs were individually placed before each PBS to balance optical power across the four output paths. The STDM module employed four sets of achromatic doublets (listed in Supplementary Table 2), with each set arranged in a 4F configuration and optimized with channel-specific focal lengths. Differential path lengths of 900 mm between adjacent channels introduced ~3 ns inter-pulse delays, raising the effective pulse repetition rate to 320 MHz. This value was chosen to approach the theoretical limit for single-point excitation, which is set by the fluorescence lifetime of the indicators at a ~3 ns inter-pulse interval[38], and to minimize fluorescence crosstalk. To minimize aberrations across the FOV, the EOD and multiplexing module were mounted on an inclined metal plate tilted at ~1.465° (Extended Data Fig. 1d), ensuring symmetric beam arrangement relative to the subsequent optical axis. Due to small inter-channel angles and proximity to the beam convergence point, the second lens in the first channel was trimmed into a 3.5-mm-wide strip to avoid obstructing adjacent beams (Extended Data Fig. 1e).

The output beam from the multiplexing module was first reduced in diameter by a two-lens group and reimaged onto a 36-facet polygon scanner (SA34, Lincoln Laser) rotating at 54,945 rpm (916 Hz). The imaging frame rate matched the polygon rotation frequency. The beam was then relayed to a galvanometer scanner pair (6215H, Cambridge Technology). The Galvo-X scanner performed *y*-axis scanning, with its deflection angle set by inter-channel angular differences. To improve speed and stability, Galvo-X scanner operated in a vertical bidirectional scanning mode

395 at half the frame rate (top-down for odd frames, bottom-up for even frames), minimizing flyback
396 dead time (see Supplementary Fig. 1). Galvo-Y scanner remained stationary during voltage
397 imaging due to its slower response. The last two facets of each polygon rotation were allocated to
398 accommodate Galvo-X flyback, resulting in 34 effective scanning facets per frame. It was
399 multiplied by 16 through the combined action of the EOD and STDM, yielding a total of 544 scan
400 lines covering a 524 μm vertical FOV per frame.

401 To reduce system cost, two identical TTL200MP lenses (Thorlabs) were used as the scan and
402 tube lenses. A dichroic mirror (DM_1; FF735-Di02-25x36, Semrock) transmitted excitation light to
403 the objective while reflecting fluorescence toward the detection path. A second dichroic mirror
404 (DM_2; FF552-Di02-25x36, Semrock) then split the fluorescence into two spectral channels. Each
405 channel was equipped with a photomultiplier tube (PMT; H7422-40, Hamamatsu) coupled to an
406 amplifier (C11184, Hamamatsu; DC-300 MHz), and signals were digitized at 2.56 GHz using a
407 vDAQ™-HS system (Vidrio Technologies).

408 In addition to kilohertz voltage imaging, the HS2PM system supported low-frame-rate and
409 bright-field imaging modes. For low-frame-rate scanning, the system switched to a conventional
410 two-photon mode: the polygon scanner and EOD were disabled, and only one channel in the
411 multiplexing module remained open, while the Galvo-X and Galvo-Y scanners worked in tandem.
412 This configuration provided a large FOV of 2.06 × 2.06 mm^2 (1024×1024 pixels, 1.71 Hz,
413 Supplementary Fig. 2a), enabling rapid target identification after which the system could switch
414 back to kHz voltage imaging of the region of interest (see Supplementary Video 2 for comparison).
415 The system also offered a bright-field imaging mode, where an LED source was used. Under
416 remote control, the movable mirror (MM; Extended Data Fig. 1c) could be moved into the path
417 between the scan and tube lenses, blocking the 920 nm laser and directing LED light to the

418 objective. Simultaneously, DM_1 was removed from the optical path, allowing sample-reflected
419 light to reach the eyepiece or CCD camera for bright-field visualization.

420 **Optimized fluorescence collection system**

421 The dichroic mirror split the collected fluorescence into two channels, each equipped with a
422 bandpass filter (FF01-593/46, FF01-520/35-25, Semrock). Both channels featured identical optical
423 assemblies. Fluorescence exiting the objective's back aperture first passed through a 75 mm focal
424 length plano-convex lens (LA1386-A, Thorlabs; Ø38.1 mm). The beam then traversed a second
425 plano-convex lens (LA1274-A, Thorlabs; f = 40 mm, Ø30 mm) and an aspheric condenser lens
426 (ACL25416U-A, Thorlabs; f = 16 mm, Ø25.4 mm), spaced 1 mm apart. This dual-lens
427 configuration coupled the fluorescence to a custom conical frustum light guide (front surface
428 Ø5.85 mm, back surface Ø4.45 mm). A critical innovation involved integrating the CFCL into the
429 7.4 mm pathway between the PMT's front window and its internal protective window. For
430 assembly, the 8.5 mm long CFCL was coaxially bonded to a glass disc (Ø14 mm, 1 mm thickness).
431 An annular spacer was first placed at the original PMT window position to create a buffer zone.
432 The CFCL-disc assembly was then inserted into the pathway, with UV-curable adhesive used to
433 co-axially bond the PMT housing, spacer, and glass disc.

434 We performed Zemax non-sequential ray-tracing simulations of fluorescence emission at
435 varying angles from the objective's back aperture. The simulations utilized an Embedded Monte
436 Carlo approach, tracing 5×10^5 rays per angle at a wavelength of 550 nm. Collection efficiency at
437 each emission angle was determined as the ratio of optical power recorded by the detector to the
438 1-W input power (Extended Data Fig. 3).

439 **Fluorescence crosstalk characterization**

To quantify temporal and spatial crosstalk in scattering tissue, we imaged a cranial-window-implanted mouse expressing ASAP5. For each measurement, only one scanning region was excited at ~50 mW by enabling a single channel of the multiplexing module, while fluorescence was collected from all four channels. This procedure was repeated sequentially for each of the four channels. At each imaging depth from 150 to 700 μm below the pial surface (in 50 μm steps), 10,000 frames were acquired at 916 Hz, and average intensity images were generated (see Extended Data Fig. 7a). It should be noted that while the schematics in Fig. 1 and Extended Data Fig. 2 illustrate four scanning regions (A-D) for conceptual clarity, the actual spatiotemporal arrangement of the four channels after the multiplexing module is shown in Extended Data Fig. 7a, where channels 1-4 represent the true scanning sequence and spatial layout. Pixels within the excited subregion corresponding to ASAP5 fluorescence were defined as those whose intensities ranked among the top 5%. To quantify crosstalk, the mean intensity of these bright pixels in the excited subregion was compared to the mean intensity of spatially corresponding pixels in each non-excited subregion. The crosstalk percentage was calculated as the ratio of the non-excited to excited mean intensities (see Extended Data Fig. 7b).

Synchronization and data acquisition

A complete synchronization diagram of the scanning and acquisition system is provided in Extended Data Fig. 8. Polygon scanner synchronization was achieved using a 780-nm reference laser (LM15-780, Lbtek) focused via a plano-convex lens (LA1484-A, Thorlabs) onto the polygon mirror (Extended Data Fig. 1f). The reflected beam was detected by a photodiode (S2506-02, Hamamatsu) mounted on a dedicated circuit board (3-1-5248-900-00, Cambridge Technology), which generated a TTL start-of-line (SOL) pulse at each polygon facet transition. This SOL pulse triggered the data acquisition system, ensuring temporal alignment between the scanning trajectory

463 and signal digitization. The laser clock was synchronized with the SOL signal, galvanometer
464 drivers, AWG, and acquisition card to generate raw data frames that were subsequently
465 reconstructed into spatially accurate images.

466 The high-speed acquisition system could sample each laser pulse 32 times, but the data
467 acquisition setting was configured differently for voltage imaging and FLIM modes. In voltage
468 imaging, acquisition windows for the four multiplexed channels were strategically aligned to
469 capture each channel's peak signal. Each window spanned 4 consecutive sampling points within
470 the 32-point pulse cycle, separated by 4-point intervals to minimize inter-channel crosstalk
471 (Extended Data Fig. 4a). This configuration enabled full-FOV imaging at 916 Hz. In FLIM mode,
472 the system was reconfigured to estimate fluorescence lifetimes *in vivo* by blocking two non-
473 adjacent multiplexing channels (e.g., channels 2 and 4), effectively imaging half of the full FOV.
474 The remaining two active channels operated with a 6.25 ns pulse interval. Each pixel was acquired
475 by two sampling windows: one capturing the prompt intensity and the other capturing the intensity
476 after a 1.5625 ns delay (Fig. 1i; Extended Data Fig. 4b). The intensity ratio between these windows
477 was then used for pixel-wise lifetime calculation.

478 **FLIM**

479 Given that fluorescence emission follows an exponential decay, we use τ to represent fluorescence
480 lifetime, and the fluorescence intensity F(t) can be expressed as $F_0 e^{-t/\tau}$. Fluorescence was
481 simultaneously sampled by two contiguous, width-matched detection windows, producing two
482 subareas' images corresponding to the prompt (P) and delayed (D) intensity:

483
$$\begin{cases} F_{\mathrm{P}} = \int_0^T F(t)dt \\ F_{\mathrm{D}} = \int_T^{2T} F(t)dt \end{cases} \tag{1}$$

484 where T presents the equal-time ~1.5625 ns during each sampling window. The intensity ratio ξ
485 between the two adjacent sampling windows directly correlates with the fluorescence lifetime:

486 $$\xi = F_{\mathrm{D}} / F_{\mathrm{P}} = e^{-T/\tau} \quad (2)$$

487 Therefore, the two-subarea pair encoded nanosecond time information in their intensity ratio. The
488 relationship curve between ξ and τ is shown in Fig.1i. We can compute the lifetime values for each
489 pixel using this intensity ratio, and the principle is similar to existing wide-field lifetime imaging
490 techniques[39].

491 **Animal preparation**

492 All mice used were male and of a C57BL/6JNifdc background, purchased from Charles River
493 Co., Ltd. All mice were socially housed in the Animal Facility at the Department of Laboratory
494 Animal Science at Fudan University under a 12 h light/dark cycle (lights on 08:00, off 20:00) and
495 room temperature with food and water ad libitum. All animal experiments were conducted in
496 accordance with the guidelines of the Institutional Animal Care and Use Committee of the
497 Department of Laboratory Animal Science at Fudan University (2022JSITBR-013, 2024-JSYY-
498 035). *In vivo* voltage imaging data in this study were collected from 9 wild-type mice, six eight-
499 week-old male mice were injected with pAAV-CaMKiia-ASAP5-Kv (a generous gift from
500 Michael Z. Lin in Stanford University) in the sensory cortex (2/1.5 mm left medial-lateral and -
501 1.06 mm anterior-posterior to lambda), and the other three were injected with AAV-CaMKii-JEDI-
502 2P-Kv in the same site at the same age. The microinjection pump (KDS 130, KD Scientific) was
503 used to inject at a rate of 50 nl/min at each viral injection site, with a total volume of 200 nl using
504 a 10 μl Hamilton syringe (Model 701N). For visualization of GEVI-labeled neurons via
505 fluorescence microscopy, the mice underwent cranial window implantation surgery after 2 months
506 for sufficient expression of the indicators. In brief, mice aged 4 months were anesthetized with

507 1~2% isoflurane (RWD Life 524 Science #R510-22-10) and head-fixed in a stereotaxic apparatus,
508 then a 4 mm circular craniotomy was made using a high-speed dental drill over the virus injection
509 position. A glass window made of a single coverslip customized by us was flushed with the skull,
510 and sealed by a tissue adhesive (VetBond, 3M). A stainless-steel head- fixator was then firmly
511 attached to the skull with dental acrylic and cyanoacrylate adhesive (ergo, 5800). Implanted mice
512 were allowed a one-week recovery period, during which they received daily intraperitoneal
513 injection of Dexamethasone (ST1254, Beyotime Biotechnology) and Cefazolin (Shanghai Macklin
514 Biochemical Technology #HY-B0712A). Both drugs were prepared as 0.5 mg/ml solutions, and
515 each mouse was injected 0.1 ml of each drug. Two-photon calcium imaging was conducted in a 6-
516 month-old male transgenic CaMKii-Cre::Ai162 mice derived from the crossbreeding of CaMKii-
517 Cre (#C001015, from Cyagen Biosciences) and Ai162 mice (#031562, from Jackson Labs). All
518 surgical procedures and pharmacological manipulations followed the same protocol used for the
519 voltage indicator-labelled mice, except that a 6-mm-diameter craniotomy was made over the
520 central brain region. Mice were then adapted to head fixation for 30 mins per day over one week
521 before imaging. After imaging experiments, unless specified, experimental animals were
522 euthanized by CO_2 asphyxiation followed by cervical dislocation.

523 ***In vivo* voltage imaging**

524 Voltage imaging was performed on awake, head-fixed mice using a total excitation power of
525 200 mW (measured post-objective). To characterize the imaging depth of HS2PM, spontaneous
526 neuronal activity was recorded at 50 μm intervals starting from 150 μm below the pial surface until
527 the signal could no longer be clearly resolved. At each depth, 10-second movie was acquired at
528 916 Hz for subsequent analysis of depth-dependent signal characteristics.

For sensory-evoked voltage recordings, two distinct whisker stimulation paradigms were applied contralateral to the imaged hemisphere. The first paradigm consisted of alternating periods of 3 s baseline, 3 s air-puff stimulation, and another 3 s post-stimulation, forming a single 9-s imaging trial. Each mouse was imaged across three different FOVs, with 10 trials recorded per FOV. The second paradigm featured transient whisker stimulation during 20-s continuous imaging sessions, using repeated 1-s air puffs with 1-s inter-stimulus intervals. For this paradigm, five trials were acquired per FOV across three FOVs per mouse. All stimuli were delivered at ~30 psi through a 20-gauge blunt needle positioned 1-2 cm from the mouse's face.

Image processing and analysis

The HS2PM image processing workflow consists of the following steps: raw data rearrangement, crosstalk removal, motion correction, denoising and extraction of neuronal voltage traces (Extended Data Fig. 5). Each raw data frame initially contained 2284 × 34 pixels. Each raw frame represented one of four multiplexed scanning regions within the full FOV. To reconstruct complete frames, four consecutive raw frames were combined into a 571 × 544 composite image at the system's native 916 Hz acquisition rate. Since the system employed bidirectional Galvo-X scanning—with opposite scanning directions for odd and even frames—a correction algorithm was applied during frame rearrangement to eliminate jitter artifacts caused by the alternating scan orientations. A custom algorithm was developed to remove crosstalk arising from the STDM scheme. This crosstalk was observed between temporally adjacent channels and negligible between non-adjacent channels. The relationship between the measured signal in each channel (CH_i, i=1, 2, 3, 4) and the underlying crosstalk-free signal (P_j, j=1, 2, 3, 4) was modeled as a linear mixing process:

551 $$\begin{cases} \mathrm{CH}_1 = \mathrm{P}_1 + \alpha_4 \mathrm{P}_4 \\ \mathrm{CH}_2 = \mathrm{P}_2 + \alpha_1 \mathrm{P}_1 \\ \mathrm{CH}_3 = \mathrm{P}_3 + \alpha_2 \mathrm{P}_2 \\ \mathrm{CH}_4 = \mathrm{P}_4 + \alpha_3 \mathrm{P}_3 \end{cases} \quad (3)$$

552 in which α_i is the crosstalk ratio of each temporal adjacent channel. The demixed crosstalk-free
553 frame was obtained by solving the equations, and the images before and after crosstalk demixing
554 process were presented in Extended Data Fig. 5c and 5d for comparison. The original crosstalk
555 ratios were determined by the measured crosstalk in Extended Data Fig. 7, and then optimized
556 during the demixing process. Neurons were identified and selected manually on the averaged
557 images. Finally, the Volpy algorithm was used for motion correction, denoising and extracting
558 neuronal voltage activity traces.

559 For voltage imaging data analyses, $-\Delta F/F$ was calculated as the negative fractional change in
560 fluorescence relative to the baseline level. The fluorescence baseline was obtained by applying a
561 rolling percentile filter (50%, 500-ms window) to the mean-intensity trace of the ROI. We
562 classified optically identified action potentials (APs) as voltage transients only if they were labeled
563 as spikes by the VolPy algorithm and their peak $-\Delta F/F$ exceeded 2.5 times the standard deviation
564 of the voltage trace within a 20-frame (~21.84 ms) rolling window centered on the spike.
565 Subthreshold $-\Delta F/F$ was calculated through the application of a 20-Hz, 5th-order Butterworth
566 lowpass filter to the $-\Delta F/F$ voltage trace. Spike rate traces were obtained by applying a rolling 20-
567 ms window on the $-\Delta F/F$ voltage traces and calculating the spike rate within this window.
568 Alongside the $-\Delta F/F$ traces, the SNR traces were calculated as the ratio of ΔF (the functional
569 change in fluorescence) to $\sqrt{F}$ (Poisson noise). The neuronal spike SNR was determined by
570 averaging the SNR values at the time points of identified spikes. Photobleaching rate was
571 quantified as the slope of a linear fit to the normalized fluorescence decay relative to its initial

572 intensity. For sensory-evoked responses, neuronal latency was defined as the interval from
573 stimulus onset to the first subsequent action potential. For calcium indicator-labeled neurons, the
574 baseline fluorescence was calculated as the mean intensity of the initial 30% of each recording
575 trace. All statistical analyses were conducted using MATLAB (R2022b, MathWorks) or built-in
576 tests in SPSS (v20, IBM).

577 **Photobleaching**

578 Photobleaching measurements were conducted using the HS2PM on male adult mice with
579 implanted cranial windows. The tissue was scanned and imaged at a frame rate of 916 Hz using a
580 total laser power of 200 mW at the sample, and the scanning imaging was performed continuously
581 for 60 min. ROIs were extracted manually. Each data point in Fig. 2l corresponded to the mean
582 fluorescence intensity of every 50th frame across one session of imaging. Photobleaching rates
583 were determined by normalizing all data points to the first data point.

584 **Photodamage**

585 Three adult mice injected with virus and implanted with cranial windows were used for the
586 experiments. The laser exposure location was selected away from the viral injection sites: we first
587 found labeled neurons by using the microscope and then moved the stage until no neurons were
588 found in the FOV. For each photodamage session, continuous laser scanning was performed for 1
589 hour at a total power of 200 mW. Twenty hours after laser exposure, mice were anesthetized with
590 1% pentobarbital sodium and transcardially perfused with heparinized saline (20 U/mL), followed
591 by 4% paraformaldehyde (PFA). Cranial windows were removed to expose the cortex. Brains were
592 postfixed in 4% PFA at 4 °C overnight, then sequentially cryoprotected in 20% and 30%
593 sucrose/PBS solutions, and sectioned into slices with the thickness of 25 μm using a vibratome.

594 To assess photodamage after 1 h of continuous imaging through immunostaining, brain slices were
595 first incubated in blocking solution (5% normal donkey serum/PBST, where PBST was PBS
596 supplemented with 0.3% Triton X-100) for 1 hour. Slices were then incubated with the primary
597 antibody (monoclonal antibody against HSP70/HSP72, Enzo Life Sciences #ADI-SPA-810,
598 diluted at 1:400) at 4 °C overnight, followed by labeling with the secondary antibody (Alexa
599 Fluor® 594 AffiniPure® Donkey Anti-Mouse IgG (H+L), Jackson #715-585-151) for 2 hours at
600 room temperature. Slices were mounted with Antifade Mounting Medium containing DAPI
601 (Beyotime Biotechnology #P0131). Olympus SLIDEVIEW VS200 was used to image the slices
602 with Olympus OlyVIA software. Photodamage was assessed by comparing antibody labeling
603 intensity in the laser-exposed region and its corresponding area in the contralateral hemisphere.
604 Fluorescence intensity was measured in three cortical regions per slice: the laser exposure site, a
605 contralateral control area at least 1 mm away from the laser exposure site (to account for potential
606 effects of chronic window implantation or virus injection unrelated to laser exposure), and the
607 surgically injured region for chronic window implantation. For each region, relative fluorescence
608 was calculated by taking the ratio of its average fluorescent intensity to that of the control area.
609 Data analysis was performed using Welch's t test (n = 3 mice).

610 **Neural decoding and clustering**

611 To assess how population activity encodes air puff stimuli, we trained linear SVM decoders[44] on
612 two representations: population spike-rate vectors and subthreshold $\Delta F/F$ traces. Data were
613 collected from 3 mice (9 FOVs, 5 trials per FOV). For each trial, we extracted neural activity from
614 1 s before to 1 s after stimulus onset and assigned binary labels (1 = puff, 0 = no puff). Within each
615 trial we slid a 20-ms window in 10-ms steps. In every time window, we evaluated the average
616 activity of each neuron at each trial. We used the standard LIBSVM toolkit[45] and fed the SVM

617 module with neural activity and binary label. Training and testing were performed using a 10-fold
618 cross-validation procedure. Data were divided into 10 equally sized and disjoint sets; in each fold,
619 one of the sets (10% of the data) was used for testing, and the remaining 9 sets (90% of the data)
620 were used for training the classifier. In each fold, the SVM regularization parameter was
621 determined using a grid-search optimization using another (internal) 10-fold partition of the
622 training data. Overall, the 10-fold cross-validation testing process resulted in 10 accuracy rates per
623 time bin.

624 For k-means clustering in Figs. 3q and 3r, we divided the temporal traces into two stages
625 (during air puff stimulus and without stimulus), and used the standard deviations of spike rate and
626 subthreshold electrical activity during corresponding stages as features for each neuron. PCA
627 method is first used to reduce the feature dimensionality from 4 to 3. We then used k-means method
628 to separate neurons into four different clusters based on the features, after which the clustered
629 groups of neurons can be precisely distinguished.

630 ***In vivo* vascular imaging and blood flow velocity measurement**

631 Two wild-type mice were used for vascular imaging. Both were briefly anesthetized with
632 isoflurane before retro-orbital injection with fluorescent solutions (200 μL each, 5% w/v): one
633 mouse received 70-kDa dextran-conjugated Texas Red, and the other was injected with 70-kDa
634 dextran-conjugated Rhodamine B. The mice were then head-fixed under the objective lens, and
635 kilohertz intensity images were recorded for blood flow velocity measurement, while half-FOV
636 images were recorded for calculating and comparing different fluorescence lifetimes.

637 Structural images were acquired for 5 s using the TPVI system at a frame rate of 916 Hz and
638 a post-objective power of 200 mW via a 25× Olympus objective. First, to better analyze the
639 temporal changes in blood flow and fluorescence lifetime, the raw vascular imaging data were

640 denoised using the SUPPORT algorithm[46] to eliminate random shot noise. The denoising model
641 was trained using the raw data acquired by HS2PM, while its hyperparameters adopt the default
642 parameters of the SUPPORT algorithm. Vessel segments were then manually traced over the
643 average intensity of the recorded structural images (150 μm below pial surface), and the
644 kymographs of the vessels were plotted for blood flow velocity calculation. Kymographs were
645 generated by plotting pixels along the traced blood vessels (horizontal axis) over time (vertical
646 axis), with dark diagonal streaks resulting from red blood cells (RBCs) moving along the vessel
647 segment. An automated method developed by Chhatbar and Kara was used to measure the slope
648 of RBC streaks using Sobel filtering and iterative Radon transforms[47], and then the blood flow
649 velocity of each segment was calculated every 92 frames (~0.1 s). The diameter of the vessel
650 segment was measured by extracting the vessel boundary: pixel brightness along a line orthogonal
651 to the vessel's length was extracted, and the diameter was determined as the distance between the
652 two edges detected by a Radon transform.

653 **References**

654 1. Knöpfel, T. & Song, C. Optical voltage imaging in neurons: moving from technology development to
655 practical tool. *Nat. Rev. Neurosci.* **20**, 719–727 (2019).
656 2. Lin, M. Z. & Schnitzer, M. J. Genetically encoded indicators of neuronal activity. *Nat. Neurosci.* **19**, 1142–
657 1153 (2016).
658 3. Grienberger, C., Giovannucci, A., Zeiger, W. & Portera-Cailliau, C. Two-photon calcium imaging of
659 neuronal activity. *Nat. Rev. Method. Prim.* **2**, 1–23 (2022).
660 4. Abdelfattah, A. *et al.* Neurophotonic tools for microscopic measurements and manipulation: status report.
661 *Neurophoton.* **9**, (2022).
662 5. Tokunaga, T., Sato, N., Arai, M., Nakamura, T. & Ishihara, T. Mechanism of sensory perception unveiled
663 by simultaneous measurement of membrane voltage and intracellular calcium. *Commun. Biol.* **7**, 1150
664 (2024).
665 6. Peterka, D. S., Takahashi, H. & Yuste, R. Imaging voltage in neurons. *Neuron* **69**, 9–21 (2011).
666 7. Platisa, J. *et al.* High-speed low-light *in vivo* two-photon voltage imaging of large neuronal populations.
667 *Nat. Methods* **20**, 1095–1103 (2023).

8. Villette, V. *et al.* Ultrafast two-photon imaging of a high-gain voltage indicator in awake behaving mice. *Cell* **179**, 1590-1608.e23 (2019).
9. Chen, T.-W. *et al.* Imaging neuronal voltage beyond the scattering limit. *Nat. Methods* **22**, 1366–1375 (2025).
10. Yang, H. H. *et al.* Subcellular imaging of voltage and calcium signals reveals neural processing *in vivo*. *Cell* **166**, 245–257 (2016).
11. Liu, Z. *et al.* Sustained deep-tissue voltage recording using a fast indicator evolved for two-photon microscopy. *Cell* **185**, 3408-3425.e29 (2022).
12. Wu, J. *et al.* Kilohertz two-photon fluorescence microscopy imaging of neural activity *in vivo*. *Nat. Methods* **17**, 287–290 (2020).
13. Zhang, T. *et al.* Kilohertz two-photon brain imaging in awake mice. *Nat. Methods* **16**, 1119–1122 (2019).
14. Kazemipour, A. *et al.* Kilohertz frame-rate two-photon tomography. *Nat. Methods* **16**, 778–786 (2019).
15. Truong, T. V., Supatto, W., Koos, D. S., Choi, J. M. & Fraser, S. E. Deep and fast live imaging with two-photon scanned light-sheet microscopy. *Nat. Methods* **8**, 757–760 (2011).
16. Kumar, M., Kishore, S., Nasenbeny, J., McLean, D. L. & Kozorovitskiy, Y. Integrated one- and two-photon scanned oblique plane illumination (SOPi) microscopy for rapid volumetric imaging. *Opt. Express* **26**, 13027–13041 (2018).
17. Sofroniew, N. J., Flickinger, D., King, J. & Svoboda, K. A large field of view two-photon mesoscope with subcellular resolution for *in vivo* imaging. *eLife* **5**, e14472 (2016).
18. Niesner, R., Andresen, V., Neumann, J., Spiecker, H. & Gunzer, M. The power of single and multibeam two-photon microscopy for high-resolution and high-speed deep tissue and intravital imaging. *Biophys. J.* **93**, 2519–2529 (2007).
19. Zhong, J. *et al.* FACED 2.0 enables large-scale voltage and calcium imaging *in vivo*. Preprint at https://doi.org/10.1101/2025.03.06.641784 (2025).
20. Yu, C.-H., Yu, Y., Canzano, J. S., Fei, Y. & Smith, S. L. Non-inertial scan angle multiplier for expanded fields-of-view. Preprint at https://doi.org/10.1101/2025.06.13.659647(2025)
21. Bumstead, J. R. *et al.* Designing a large field-of-view two-photon microscope using optical invariant analysis. *Neurophotonics* **5**, 025001 (2018).
22. Draelos, M., Viehland, C., McNabb, R. P., Kuo, A. N. & Izatt, J. A. Adaptive point-scan imaging beyond the frame rate–resolution limit with scene-reactive scan trajectories. *Optica* **9**, 1276–1288 (2022).
23. Podgorski, K. & Ranganathan, G. Brain heating induced by near-infrared lasers during multiphoton microscopy. *J. Neurophysiol.* **116**, 1012–1023 (2016).
24. Xiao, S., Giblin, J. T., Boas, D. A. & Mertz, J. High-throughput deep tissue two-photon microscopy at kilohertz frame rates. *Optica* **10**, 763 (2023).
25. Lin, J., Cheng, Z., Yang, G. & Cui, M. Optical gearbox enabled versatile multiscale high-throughput multiphoton functional imaging. *Nat. Commun.* **13**, 6564 (2022).

26. Zhang, Z. *et al.* High-Speed Neural Imaging with Multiplexed Miniature Two-Photon Microscopy (M-MINI2P). Preprint at https://doi.org/10.1101/2025.03.04.641573 (2025).
27. Zhao, S. *et al.* Deep two-photon voltage imaging with adaptive excitation. Preprint at https://doi.org/10.21203/rs.3.rs-5434919/v1 (2024).
28. Xiao, S. *et al.* Large-scale deep tissue voltage imaging with targeted-illumination confocal microscopy. *Nat. Methods* **21**, 1094–1102 (2024).
29. Cai, C. *et al.* VolPy: Automated and scalable analysis pipelines for voltage imaging datasets. *PLOS Comput. Biol.* **17**, e1008806 (2021).
30. Ito, J. *et al.* Whisker barrel cortex delta oscillations and gamma power in the awake mouse are linked to respiration. *Nat. Commun.* **5**, 3572 (2014).
31. Wilt, B. A., Fitzgerald, J. E. & Schnitzer, M. J. Photon shot noise limits on optical detection of neuronal spikes and estimation of spike timing. *Biophys. J.* **104**, 51–62 (2013).
32. Ebrahimi, S. *et al.* Emergent reliability in sensory cortical coding and inter-area communication. *Nature* **605**, 713–721 (2022).
33. Whitmire, C. J. & Stanley, G. B. Rapid sensory adaptation redux: a circuit perspective. *Neuron* **92**, 298–315 (2016).
34. Meng, G. *et al.* Ultrafast two-photon fluorescence imaging of cerebral blood circulation in the mouse brain *in vivo*. *Proc. Natl. Acad. Sci.* **119**, e2117346119 (2022).
35. Santisakultarm, T. P. *et al. In vivo* two-photon excited fluorescence microscopy reveals cardiac- and respiration-dependent pulsatile blood flow in cortical blood vessels in mice. *Am. J. Physiol.-Heart Circul. Physiol.* **302**, H1367–H1377 (2012).
36. Brismar, H., Trepte, O. & Ulfhake, B. Spectra and fluorescence lifetimes of lissamine rhodamine, tetramethylrhodamine isothiocyanate, texas red, and cyanine-3.18 fluorophores - influences of some environmental-factors recorded with a confocal laser-scanning microscope. *J. Histochem. Cytochem.* **43**, 699–707 (1995).
37. Kristoffersen, A. S., Erga, S. R., Hamre, B. & Frette, Ø. Testing fluorescence lifetime standards using two-photon excitation and time-domain instrumentation: rhodamine B, coumarin 6 and lucifer yellow. *J Fluoresc.* **24**, 1015–1024 (2014).
38. Brinks, D., Klein, A. J. & Cohen, A. E. Two-photon lifetime imaging of voltage indicating proteins as a probe of absolute membrane voltage. *Biophys. J.* **109**, 914–921 (2015).
39. Bowman, A. J., Huang, C., Schnitzer, M. J. & Kasevich, M. A. Wide-field fluorescence lifetime imaging of neuron spiking and subthreshold activity *in vivo*. *Science* **380**, 1270–1275 (2023).
40. Adolphs, R. The unsolved problems of neuroscience. *Trends Cogn. Sci.* **19**, 173–175 (2015).
41. Piatkevich, K. D. *et al.* Population imaging of neural activity in awake behaving mice. *Nature* **574**, 413–417 (2019).

42. Adam, Y. *et al.* Voltage imaging and optogenetics reveal behaviour-dependent changes in hippocampal dynamics. *Nature* **569**, 413–417 (2019).
43. Cheng, A., Gonçalves, J. T., Golshani, P., Arisaka, K. & Portera-Cailliau, C. Simultaneous two-photon calcium imaging at different depths with spatiotemporal multiplexing. *Nat. Methods* **8**, 139–142 (2011).
44. Boser, B. E., Guyon, I. M. & Vapnik, V. N. A training algorithm for optimal margin classifiers. In *COLT'92*: *Proc. 5th Annual Workshop on Computational Learning Theory* (ed. Haussler, D.) 144–152 (Association for Computing Machinery, 1992).
45. Chang, C.-C. & Lin, C.-J. LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.* **2**, 27:1-27:27 (2011).
46. Eom, M. *et al.* Statistically unbiased prediction enables accurate denoising of voltage imaging data. *Nat. Methods* **20**, 1581–1592 (2023).
47. Chhatbar, P. Y. & Kara, P. Improved blood velocity measurements with a hybrid image filtering and iterative Radon transform algorithm. *Front. Neurosci.* **7**, 106 (2013).

Resource Availability

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Bo Li (bo-li@sjtu.edu.cn).

Materials availability

This study did not generate new unique reagents.

Data and code availability

All data generated in this study are included in the published article. Datasets and codes are available from the corresponding author upon reasonable request.

Acknowledgements

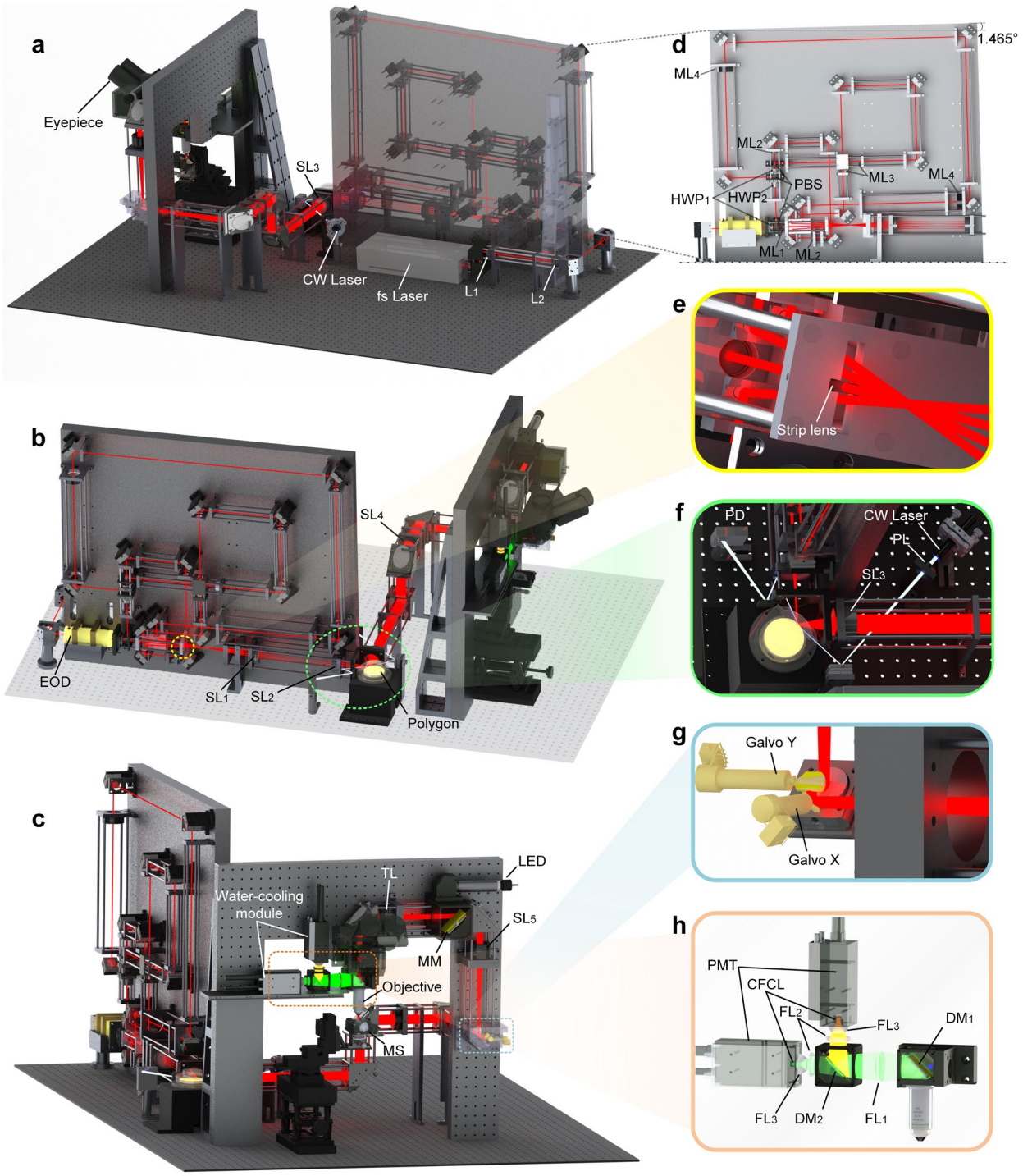
This work was supported by the National Natural Science Foundation of China (T2222006, 32471142, 32400928), China Postdoctoral Science Foundation (BX20230081), Science and Technology Commission of shanghai Municipality (20JC1419500). The authors thank Michael Z. Lin (Stanford University) for the generous gift of the ASAP5 voltage indicator.

Author Contributions

Conceptualization: B.L.; methodology: B.L., M.Z., S.P.L., C.H.K, H.Y.; biological experiments: Y.G.Z., Y.F.Z., X.Y.G., H.Y.; Zemax optical simulation: M.Z., J.H.H., F.X.; investigation: B.L., M.Z., S.P.L., H.Y.; visualization: M.Z., S.P.L.; funding acquisition: B.L., Y.M., L.C., M.Z.; project administration: B.L.; writing: B.L, M.Z., S.P.L., H.Y., Y.G.Z., Y.M., L.C., J.C.W.

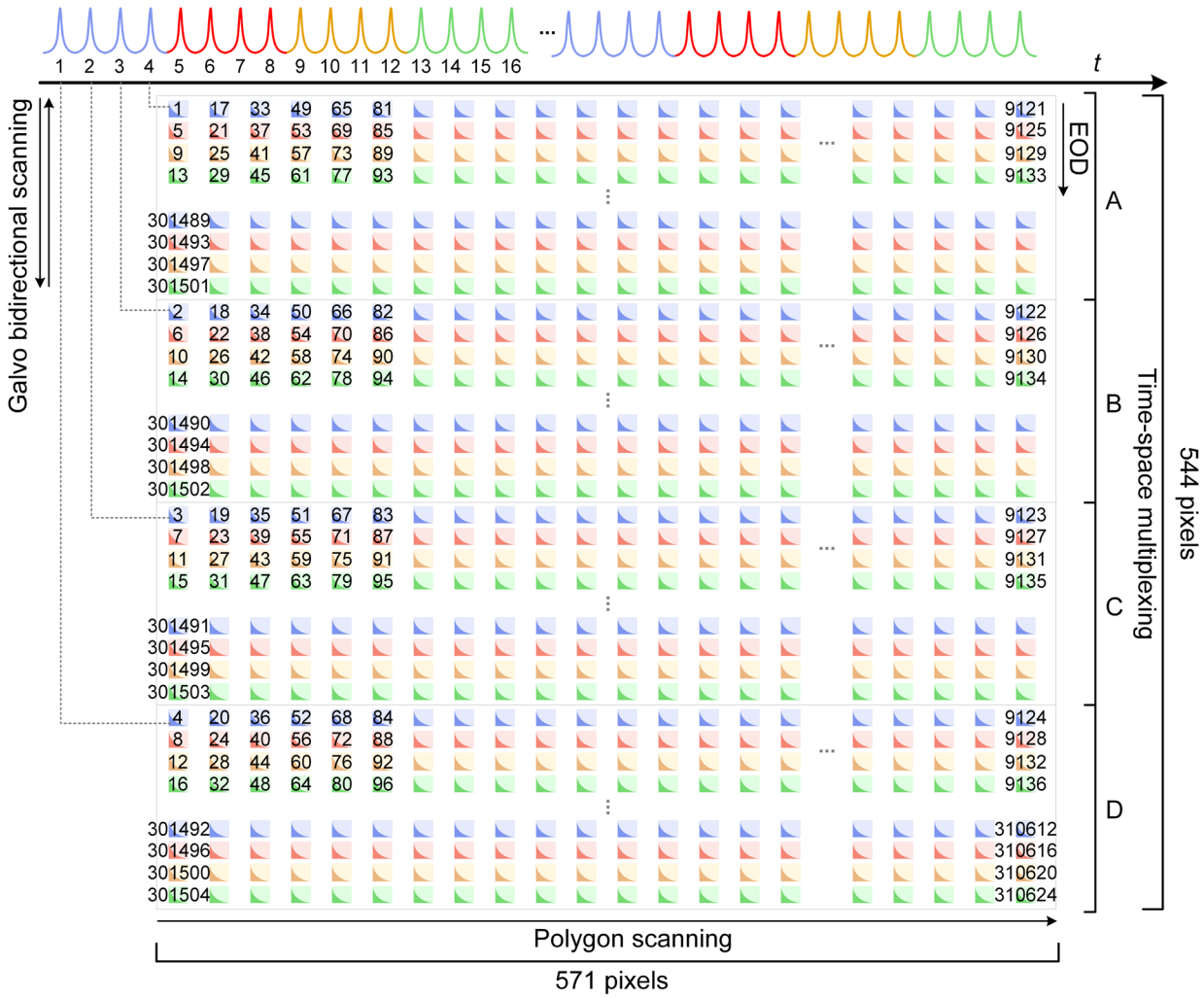
Declaration of interests

The authors declare no competing interests.

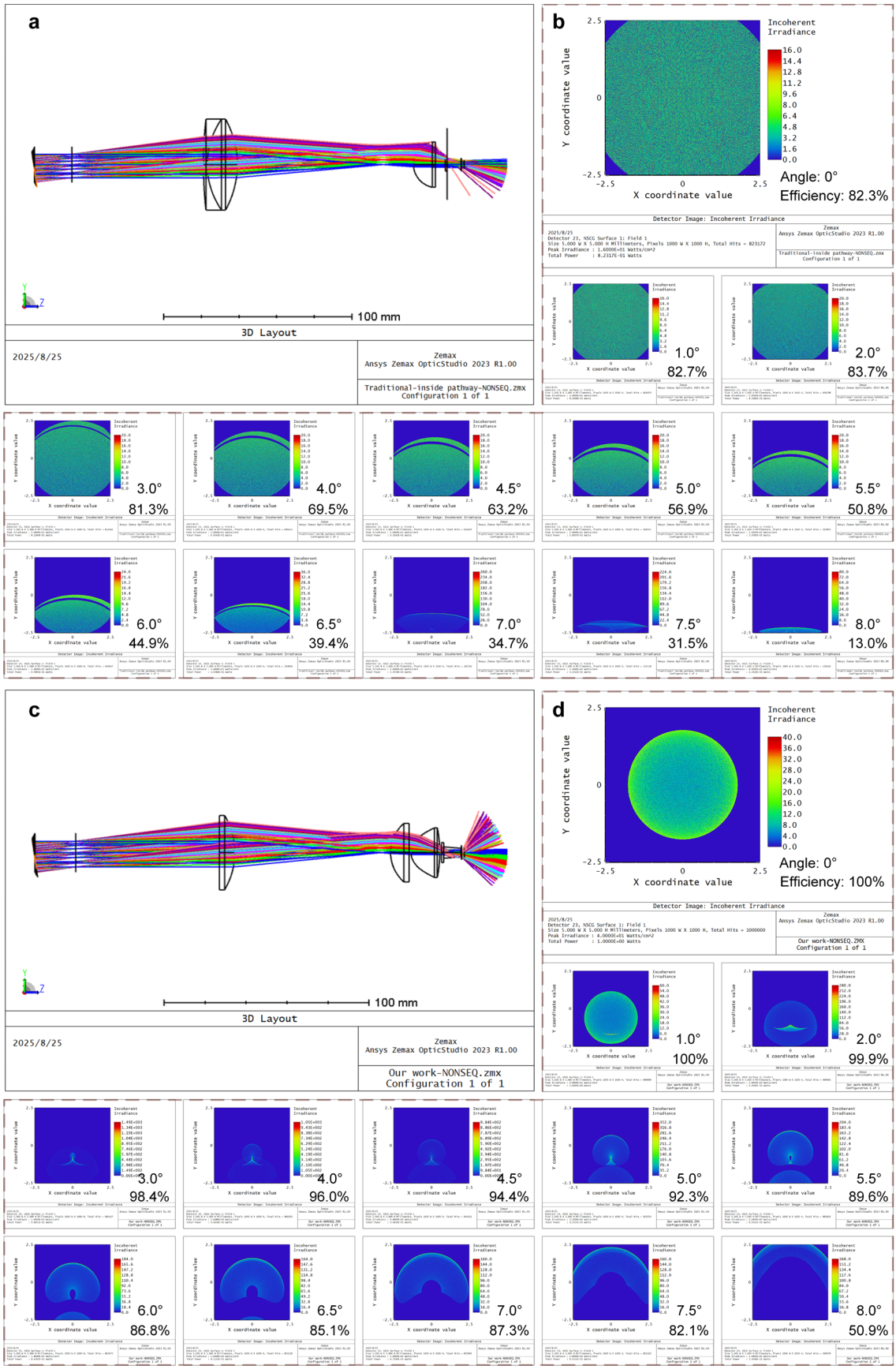


Extended Data Fig. 1 Opto-mechanical design of HS2PM. a-c, SolidWorks renderings of the complete HS2PM system, illustrating integration of optical and mechanical components. Excitation and fluorescence collection beam paths are shown. Insets highlight critical subsystems. **d**, Multiplexing module. The femtosecond laser is split into four beams using three HWPs and PBSs, with each beam routed through a distinct relay lens pair (ML_1-ML_4) and launched at a unique angle. **e**, Custom strip achromatic lens for

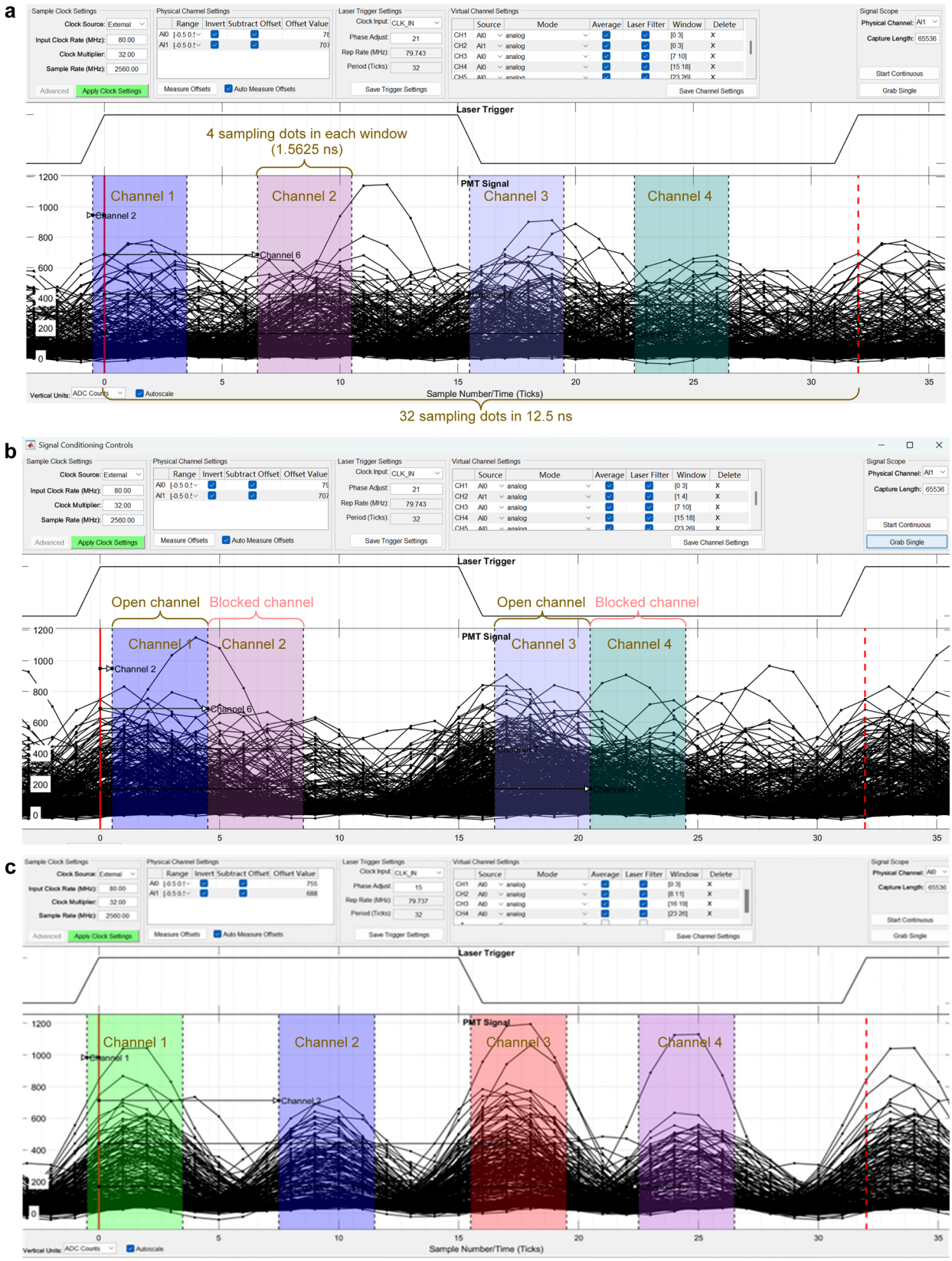
channel 1, machined on both sides to prevent obstruction of adjacent beams. **f**, Start-of-line (SOL) signal detection: a 780 nm continuous-wave laser is focused onto the polygon scanner via a plano-convex lens (PL), and reflected signals are captured by a photodiode to generate TTL sync pulses per facet transition. **g**, Galvanometer scanner assembly for bidirectional *y*-axis scanning. **h**, Dual-channel fluorescence detection system. EOD, electro-optic deflector; SL, scan lens; TL, tube lens; DM, dichroic mirror; MM, movable mirror; MS, mouse stage; LED, light emitting diode; FL, fluorescence-collection lens; CFCL, custom frustum-of-cone light guide; PMT, photomultiplier tube.



Extended Data Fig. 2 Schematic of pixel-wise order map of each frame.

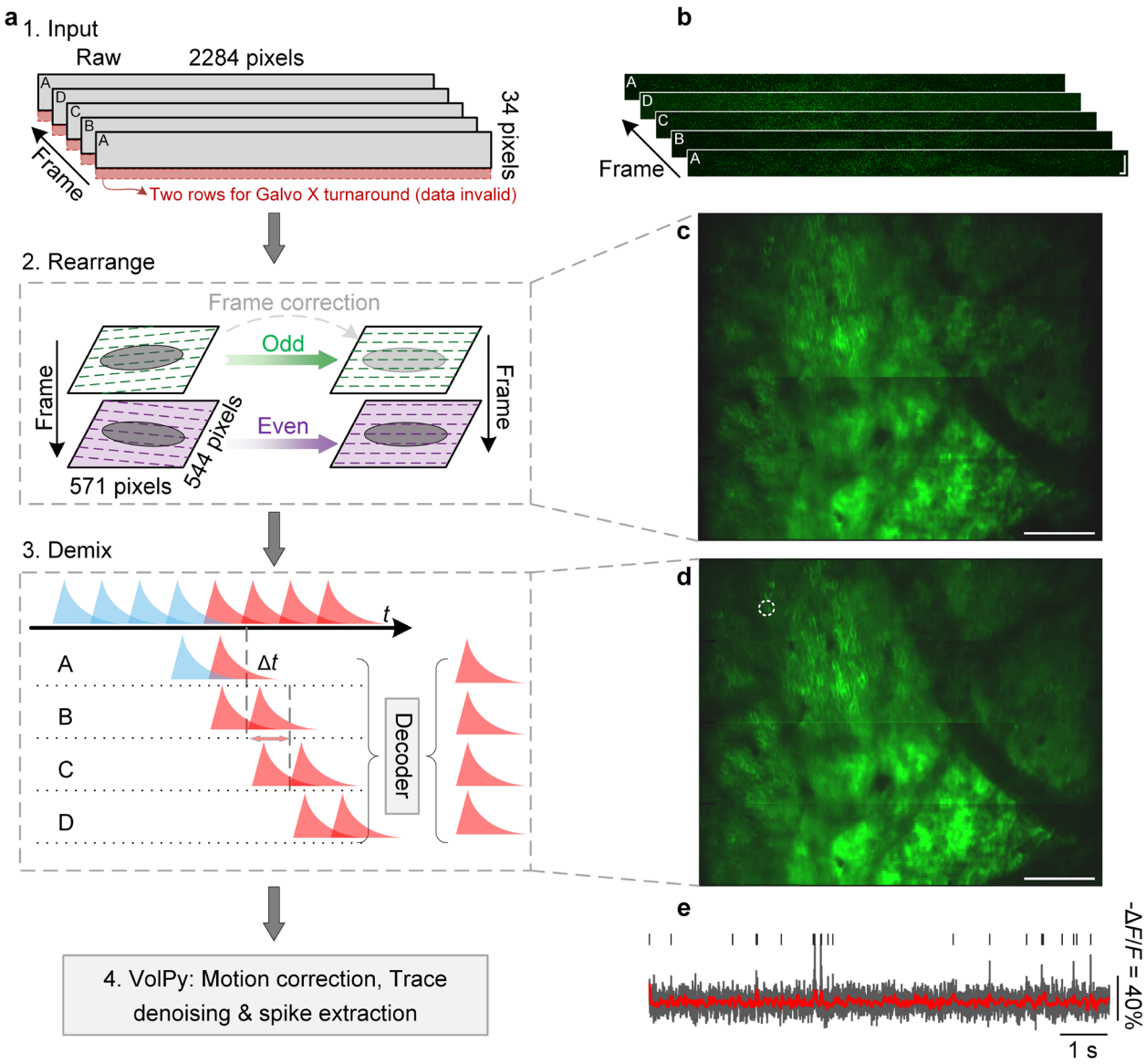

a
100 mm
3D Layout
2025/8/25
Zemax
Ansys Zemax OpticStudio 2023 R1.00
Traditional-inside pathway-NONSEQ.zmx
Configuration 1 of 1
b
Incoherent Irradiance
Y coordinate value
X coordinate value
Angle: 0°
Efficiency: 82.3%
Detector Image: Incoherent Irradiance
1.0° 82.7%
2.0° 83.7%
3.0° 81.3%
4.0° 69.5%
4.5° 63.2%
5.0° 56.9%
5.5° 50.8%
6.0° 44.9%
6.5° 39.4%
7.0° 34.7%
7.5° 31.5%
8.0° 13.0%
c
100 mm
3D Layout
2025/8/25
Zemax
Ansys Zemax OpticStudio 2023 R1.00
Our work-NONSEQ.zmx
Configuration 1 of 1
d
Incoherent Irradiance
Y coordinate value
X coordinate value
Angle: 0°
Efficiency: 100%
Detector Image: Incoherent Irradiance
1.0° 100%
2.0° 99.9%
3.0° 98.4%
4.0° 96.0%
4.5° 94.4%
5.0° 92.3%
5.5° 89.6%
6.0° 86.8%
6.5° 85.1%
7.0° 87.3%
7.5° 82.1%
8.0° 50.9%

Extended Data Fig. 3 Comparison of fluorescence collection systems between conventional and the optimized design via Zemax optical simulations. a, Schematic 3D layout of a conventional fluorescence collection system simulated in Zemax. **b**, Simulated collection efficiency of the conventional fluorescence collection system across different angles. Each angle's efficiency is defined as the ratio of power received by the detector to the initial 1 W input. **c**, 3D Zemax layout of the optimized fluorescence collection system. **d**, Simulated collection efficiency of the optimized system across different angles.

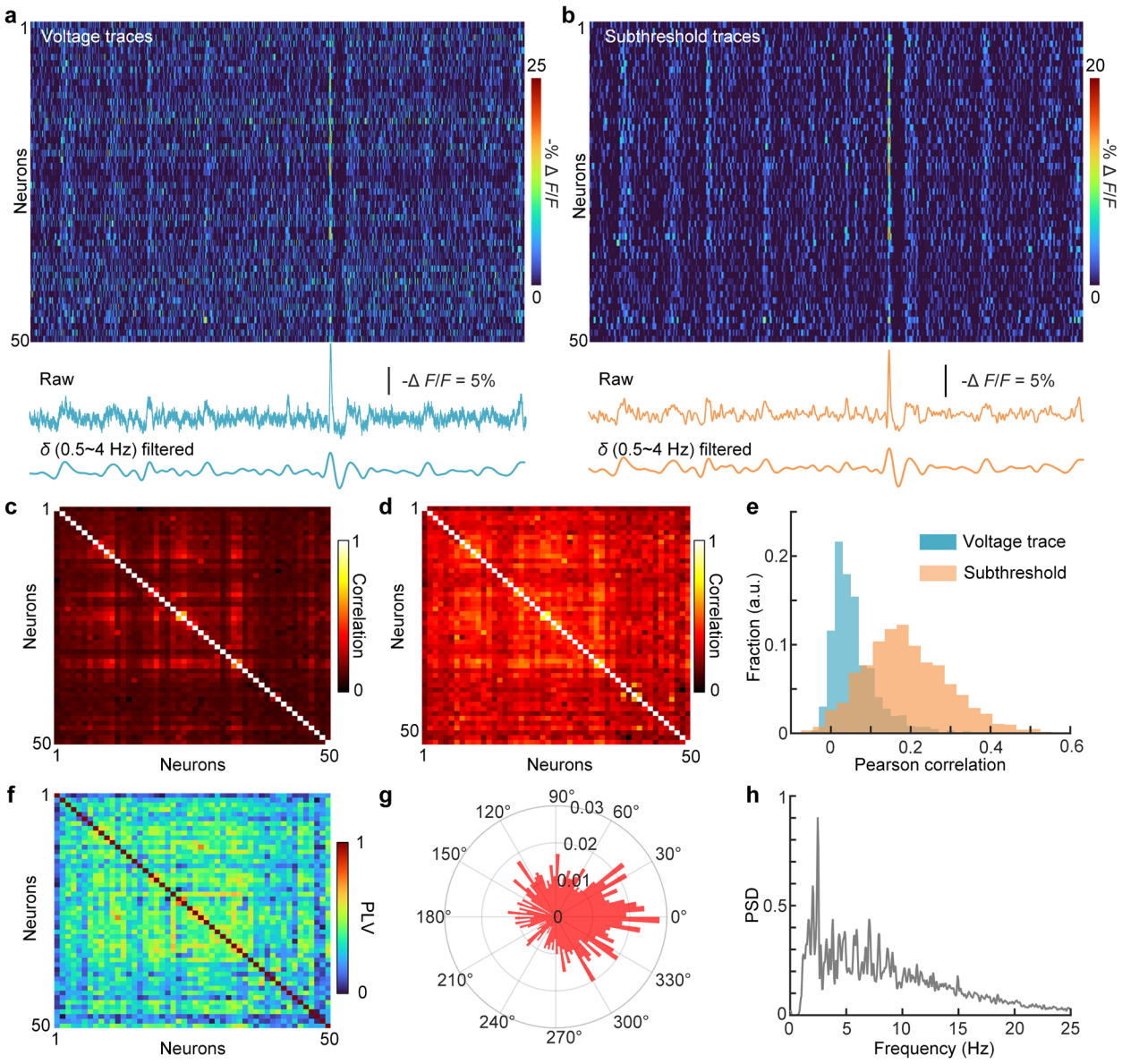

a
Laser Trigger
4 sampling dots in each window
(1.5625 ns)
PMT Signal
Channel 1
Channel 2
Channel 3
Channel 4
Sample Number/Time (Ticks)
32 sampling dots in 12.5 ns
b
Signal Conditioning Controls
Laser Trigger
Open channel
Blocked channel
Open channel
Blocked channel
PMT Signal
Channel 1
Channel 2
Channel 3
Channel 4
c
Laser Trigger
PMT Signal
Channel 1
Channel 2
Channel 3
Channel 4
Sample Number/Time (Ticks)

796 **Extended Data Fig. 4 Data acquisition settings and the collected signals of different imaging modes.**

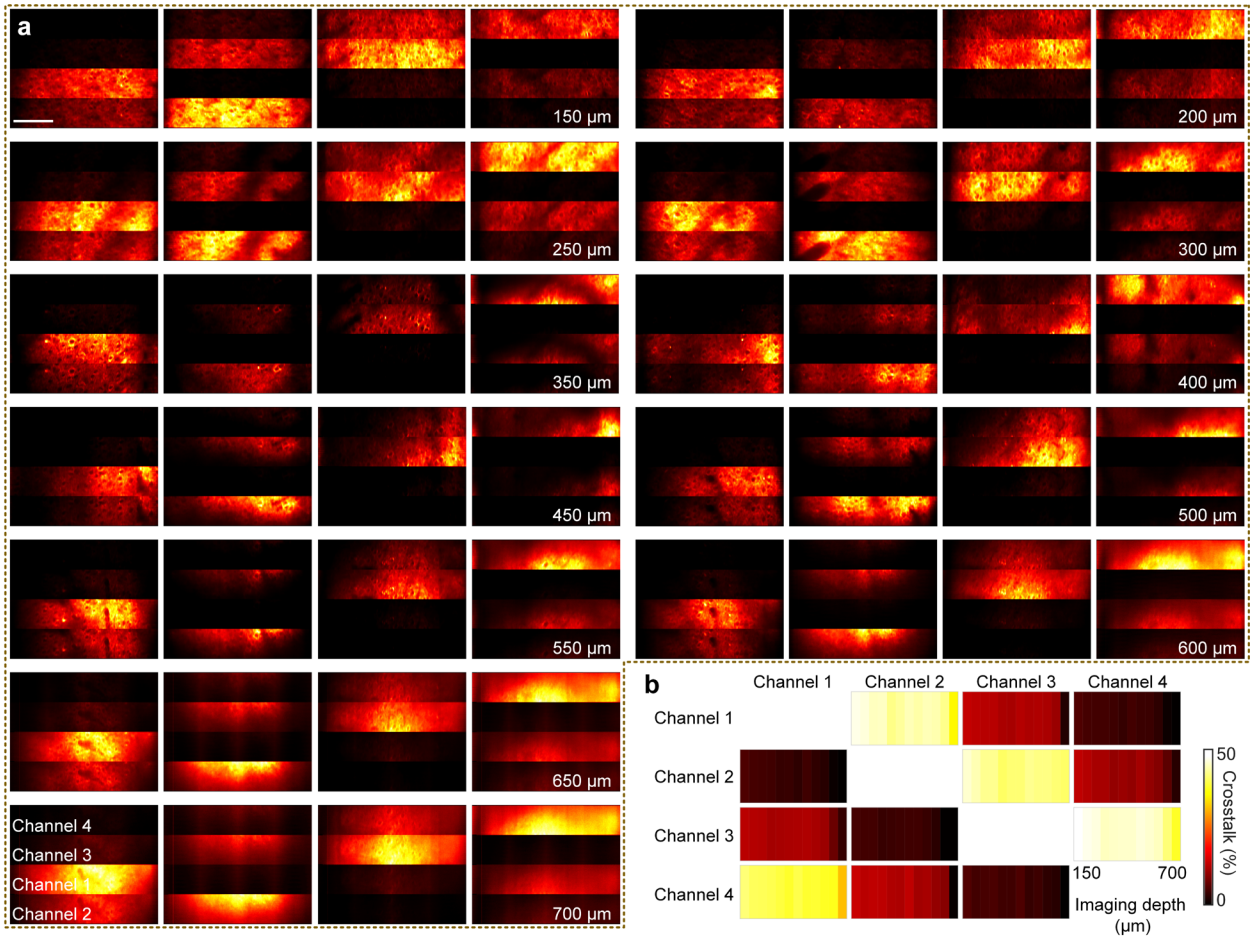
797 **a**, HS2PM, **b**, HS-FLIM, and **c**, HS2PM recorded free of cross-talk signals of *in vivo* autofluorescence.



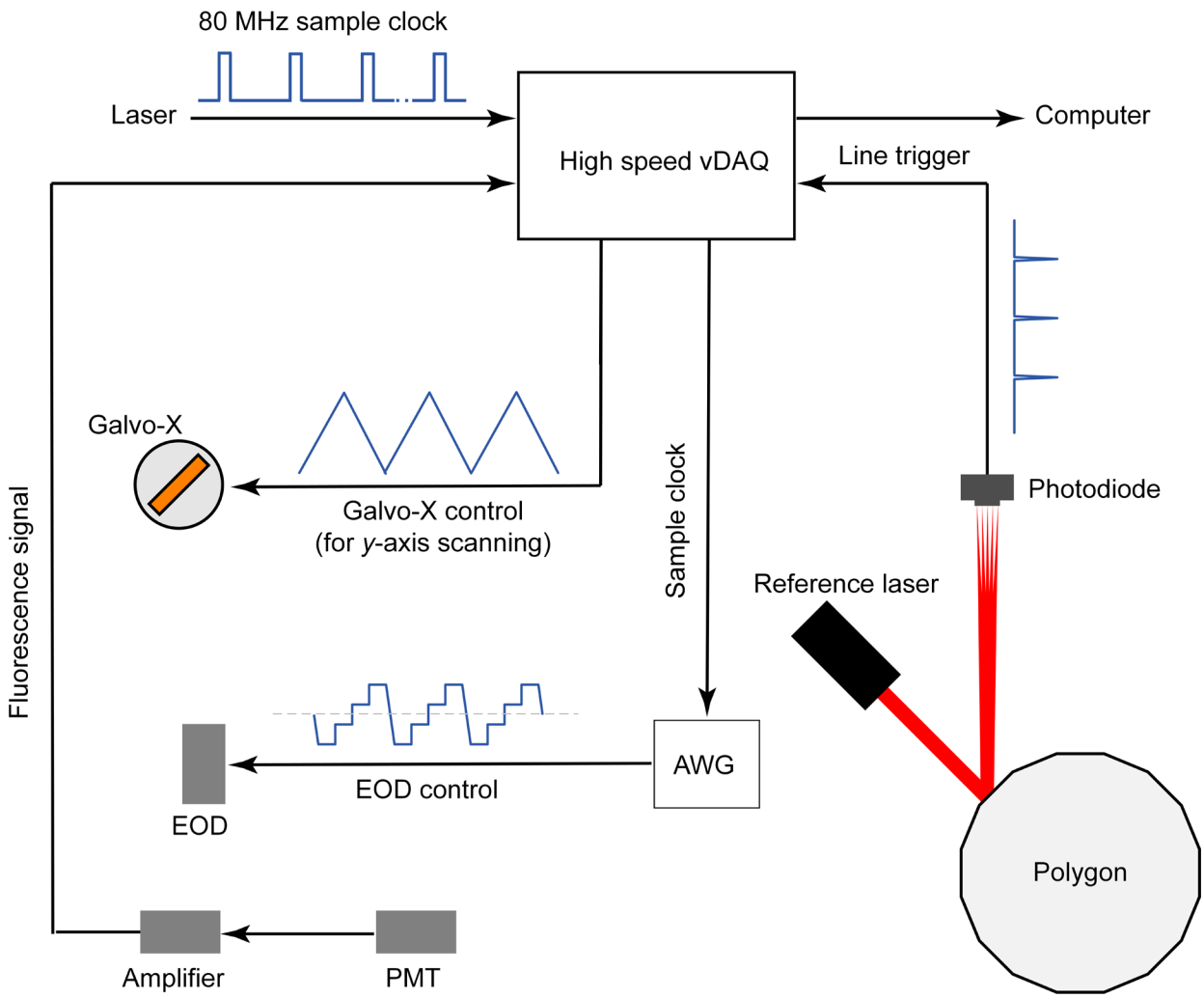
Extended Data Fig. 5 Computational pipeline for HS2PM voltage data extraction. a, Overview of the data-processing workflow for extracting neuronal voltage signals from raw imaging data. **b**, Representative raw frame (2,284 × 36 pixels). Only 34 rows are effective per frame, as 2 polygon facets are reserved for glavo turnaround during bidirectional scanning. **c**, Reconstructed raster-scanned image (571 × 544 pixels) after pixel rearrangement and odd-even frame correction. **d**, STDM demixing applied to **c** removes crosstalk across temporally multiplexed channels. **e**, Example -Δ*F*/*F* voltage trace (dark gray) extracted from a neuron identified in **d**. Red trace: subthreshold -Δ*F*/*F*; black ticks: detected spikes.



Extended Data Fig. 6 *In vivo* characterization of neuronal ensemble oscillations at 150 µm depth (related to Fig. 2f). a-b, Voltage traces and low-pass filtered subthreshold traces from 50 neurons over 10 s. From top to bottom: Heatmap of traces, averaged traces before and after filtered in the δ band (0.5-4 Hz). **c-d**, Cross-correlation matrices of voltage signals and subthreshold activities across all neurons. **e**, Distribution of Pearson's correlation coefficients for spike and subthreshold signals. **f**, Phase-locking value (PLV) matrix among neurons. **g**, Polar plots of spike phase distribution relative to the 0.5-4 Hz membrane potential oscillations across neurons. **f**, Population power spectral density (PSD) of neuronal voltages activities in **a**.



Extended Data Fig. 7 ***In vivo*** **characterization of multiplexing channel crosstalk across cortical depths.** **a**, Depth-resolved images (150-700 µm, in 50 µm steps) of four multiplexed channels. At each depth, four panels show signals detected by all four channels when only a single corresponding channel was unblocked for excitation. **b**, Crosstalk ratios among the four time-division channels as a function of depth, computed from the images shown in **a**. Scale bar: 100 µm.



Extended Data Fig. 8 Synchronization diagram of the scanning and data acquisitions system. DAQ, data acquisition; AWG, arbitrary waveform generator.