Supporting Information

"Nanofabricated Neural Probes for Dense 3-D Recordings of Brain Activity"

Gustavo Rios,[1,*] Evgueniy V. Lubenov,[1,*] Derrick Chi,[3]
Michael L. Roukes,[1,2,3,4,†] and Athanassios G. Siapas[1,2,‡]

(California Institute of Technology, Pasadena, CA 91125, USA)

[1]*Division of Biology and Biological Engineering*
[2]*Division of Engineering and Applied Science*
[3]*Division of Physics, Mathematics, and Astronomy*
[4]*Kavli Nanoscience Institute*

(Dated: September 9, 2016)

*These authors contributed equally to this work.
†Electronic address: roukes@caltech.edu
‡Electronic address: thanos@caltech.edu

Neural Probe Fabrication

Neural probes were fabricated on 200 mm (8 in) SOI wafers. The fabrication protocol consists of ∼56 processing steps requiring the use of 7 stepper field masks (field size: 22 mm × 22 mm) and both LETI's CMOS and MEMS foundries. Neural probe material composition is summarized in Table S1. Key steps used in the fabrication of the neural probes are given as follows and depicted in Figure S1:

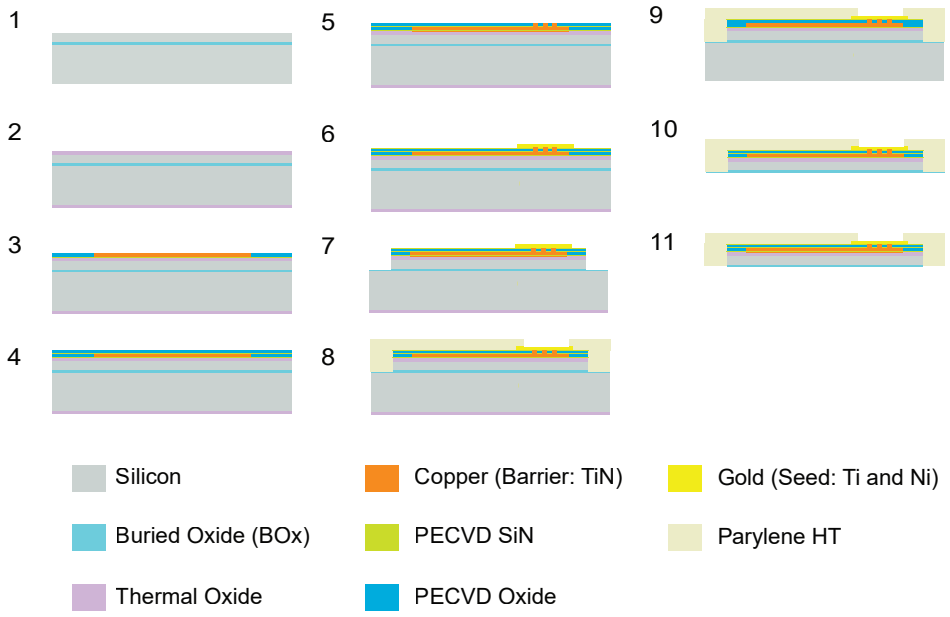

FIG. S1: **Neural probe fabrication steps. (1)** Substrate used: 200 mm SOI wafer. **(2)** Bottom insulator growth. **(3)** Copper interconnects. **(4)** Top insulator deposition. **(5)** Via etching and copper fill. **(6)** Gold microelectrode and output pad electroplating. **(7)** Top side device definition and etch. **(8)** Parylene HT deposition and etch. **(9)** Back side thermal oxide etch. **(10)** Handle layer etch. **(11)** Final device release. Color legend defines the materials used in the process. See text for details.

Bottom insulator (1-2) (1) Alignment marks are patterned and etched on the SOI wafer (device layer: 17 μm, BOx: 1 μm, handle layer: 700 μm). (2) Thermal oxide is grown. It serves as the bottom insulator for the nanoscale interconnects and as a crucial stress compensation layer. Its thickness is determined by design, as opposed to process requirements, as is the case with subsequent dielectric layers. FEM simulations

TABLE S1: **Neural probe layers and materials.**

Layer	**Material**	**Thickness**
Substrate	Si / BOx	17 μm / 900 nm
Bottom Insulator	SiO_2	0.8, 1.2, 1.5 μm
Metal Interconnects	Cu	300 nm
Top Insulator	SiO / SiN	500 nm / 40 nm
Metal Electrodes and Pads	Au	300 nm
Biocompatible Layer	Parylene HT	1.5 μm

determined that a thickness of 1.2 μm was adequate for proper stress compensation, while devices with three different SiO_2 thicknesses (0.8, 1.2, 1.5 μm) were fabricated.

Metal interconnects (3) The wafer is transferred into the CMOS foundry for this step. A dielectric layer stack of SiO / SiN (300 nm / 40 nm, SiN serves as an etch stop) is deposited using PECVD and the interconnect metal lines are patterned using a 248 nm deep ultra violet (DUV) stepper. Trenches are etched (RIE) into the dielectric stack and copper is electroplated using the CMOS industry standard damascene process (width = 300 nm, AR = 1, TiN 20 nm as barrier layer).

Top insulator (4-5) (4) A second dielectric layer stack of SiO / SiN (500 nm / 40 nm) is deposited (PECVD) on top of the copper interconnects and serves as a top insulator. SiN is used solely as an etch stop. (5) Vias are etched (RIE) through the top insulator layer at the recording electrode and interface pad locations. A second copper damascene process fills the vias. This is the last processing step in the CMOS foundry.

Metal electrodes and interface pads (6) The wafer is transferred into the MEMS foundry for the remaining procedures. Gold is electroplated (300 nm Au; 100 nm Ti and 350 nm Ni used as seed layers) and etched to define the recording electrodes and interface pads.

Top side device definition (7) Top side device definition is performed by etching through the multiple dielectric layers (RIE), the silicon device layer (DRIE) and 90% of the BOx layer (RIE, 900 nm).

Biocompatibility layer (8) Parylene HT (1.5 μm) is deposited (CVD, Specialty Coating Systems) onto the primed wafer (Silane, A174). Depositing Parylene HT after top side definition assures that the top and side walls of the shanks are fully coated. Following deposition, Parylene HT is protected with photoresist optimized for step coverage and the recording sites are patterned. Finally, Parylene HT is etched (RIE) to expose the gold recording sites.

Back side device definition (9-11) (9) The back side thermal oxide layer is patterned and etched (RIE) exposing the silicon device layer. (10,11) Final DRIE etch is performed to completely etch away the silicon from the device layer (700 μm). The left-over BOX layer (100 nm) serves as an etch stop and guarantees a smooth backside surface required for subsequent packaging steps. Finally, the thin BOX layer is etched (RIE), fully releasing the nanoprobes. Anchors keep the nanoprobes in place on the wafer until they are manually extracted.

Data Acquisition Architecture

Each module of the data acquisition (DAQ) system is designed to support 256 channels, i.e. the number of recording electrodes on a nanoprobe. Just as the nanoprobes are stacked to build a 3-D recording array, the DAQ modules are stacked to scale up the back-end data acquisition system. Figure S2 shows 4 modules assembled into a 1024-channel system. The raw analog signals from the recording electrodes feed into a custom headstage. The headstage provides signal conditioning, multiplexing, and digitization of the analog nanoprobe signals from within the implant package. Two versions of these boards were designed and built. The acute version, depicted in the block diagram, has eight 32-channel Intan chips (RHD2132), while the chronic version contains four 64-channel Intan bare dies (RHD2164). At the headstage signals are filtered (0.1 Hz–7.5 kHz passband), amplified in 2 stages (Gain = 192 V/V = 45.6 dB), multiplexed (256:8), digitized (16-bit), and transmitted out (SPI bus) through a custom cable. Custom SPI cable adapter reroutes the lines to allow interfacing directly to Intan's back-end hardware platform using its commercial SPI cables. From there signals enter an USB/FPGA interface module (Opal Kelly XEM6010). The FPGA in this module (Xilinx Spartan-6 XC6SLX150-2) is programmed to control the behavior of the headstage over SPI and to provide both data and command interface to the host computer

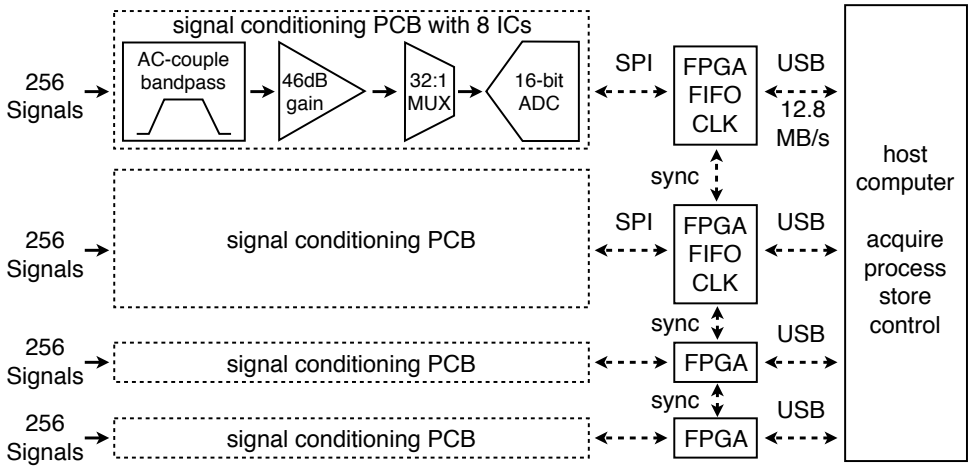

FIG. S2: **Data acquisition system block diagram.** The 1024-channel system is a stack of 4 identical 256-channel modules. Each module consists of the following components: custom headstage (signal conditioning PCB), custom 14' SPI cable tether, custom SPI cable adapter PCB, 4 Intan SPI cables, Intan evaluation board with XEM6010 FPGA Opal Kelly daughter board, USB cable. The signal path through the system is illustrated in the top headstage. Analog signals are represented by solid lines and digital signals by interrupted lines. Synchronizing pulses exchanged between the FPGA boards accomplish coordinated acquisition between the modules.

over USB 2.0 (Rhythm API). Using this platform, 256 channels can be simultaneously acquired and streamed from the headstage at sample rates up to 30 kSamples/s per channel. Since each group of 32 channels is handled in parallel through a dedicated analog multiplexer (MUX) and analog-to-digital converter (ADC), the sampling period is time divided into ∼32 conversion intervals and 8 signals are simultaneously digitized during each interval. For example, the 33.33 μs sampling period, corresponding to 30 kSamples/s channel sampling rate, is associated with ∼1 μs sampling delay between adjacent channels, i.e. the MUX switching rate and ADC conversion rate is ∼1 MHz. Intan's RHD2000 interface software and Open Ephys, both open source, were modified for compatibility with the custom headstages and enabled real-time data viewing, acquisition, and storage at the host computer. Coordinated acquisition across multiple 256-channel modules was accomplished by extending the FPGA and host computer components of the Rhythm API to support a synchronization barrier during each sample period.

Heat Dissipation Analysis

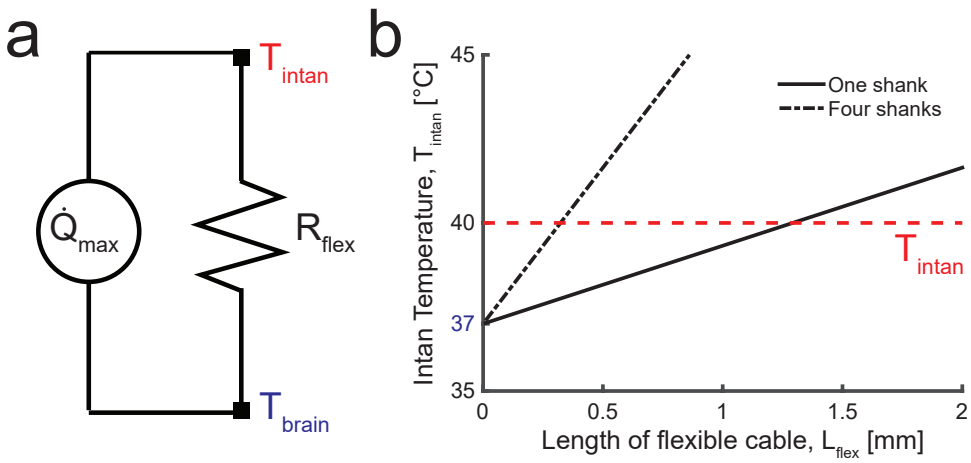

FIG. S3: **Thermal insulation by flexible cable. (a)** Simple 1-D model of heat transfer between an active headstage and a passive nanoprobe implanted in the brain represented as an equivalent circuit. $\dot{Q}_{\text{max}}$ is the maximum admissible rate of heat transfer into the brain per shank, T_{intan} is the measured steady-state temperature at the Intan signal processing IC on the headstage, $T_{\text{brain}} = 37$ °C is the temperature of the brain during normal physiological conditions and R_{flex} is the thermal resistance of the flexible cable. **(b)** Plot relating the length L_{flex} of the flexible cable to the temperature at the headstage, assuming a maximum admissible heat flux of 0.4 mW/mm^2 at the surface of one shank (solid line) and 4 shanks (interrupted line). The red dashed line indicates the temperature measured at the Intan IC.

Prolonged temperature increase in the brain as little as 0.5 °C can alter the excitability of cells, while an increase above 2 °C can have severe effects on tissue function [S1]. Therefore the heat dissipated from all active electronics, i.e. Intan ICs on the headstage, needs to be directed away from the brain. A basic guideline for implantable devices is that the heat flux through surfaces in contact with tissue should not exceed 40 mW/cm^2 (0.4 mW/mm^2) [S2, S3]. While no specific guideline has been established for brain tissue, the corresponding limit is likely lower than the figure above. Since a shank that is 5 mm long, 50 μm wide, and 20 μm thick has a surface area of $\sim$0.7 mm^2, the maximum admissible rate of heat transfer per shank is $\dot{Q}_{\text{max}} = 0.28$ mW.

To obtain a rough estimate of the heat flow between the active headstage and the animal's brain through the flexible cable we used a simple 1-D equivalent circuit model obeying

Fourier's law of thermal conduction (Fig. S3a), where the rate of heat transfer $\dot{Q}$ is proportional to the temperature difference between the headstage T_{intan} and the brain T_{brain}

$$\dot{Q} = \frac{T_{\text{intan}} - T_{\text{brain}}}{R_{\text{flex}}}$$

The coefficient of proportionality is given by the reciprocal of the thermal resistance of the flexible cable R_{flex}, which is dominated by the gold lines and can be expressed as

$$R_{\text{flex}} = \frac{L_{\text{flex}}}{k_{\text{Au}} A_{\text{Au}}}$$

where L_{flex} is the length of the flexible cable in mm, $A_{\text{Au}} = 256$ (lines) $\times$ 5 μm (width) $\times$ 0.3 μm (height) $= 3.84 \times 10^{-4}$ mm^2, and $k_{\text{Au}} = 314$ mW$\cdot$mm$^{-1}\cdot$K^{-1} is the thermal conductivity of gold.

How hot can the headstage get before the rate of heat transfer to the brain exceeds the admissible limit? Letting $\dot{Q} = \dot{Q}_{\text{max}}$ and $T_{\text{brain}} = 37$ °C in the equations above and rearranging gives

$$T_{\text{intan}} = T_{\text{brain}} + \frac{\dot{Q}_{\text{max}}}{k_{\text{Au}} A_{\text{Au}}} L_{\text{flex}} = 37 + 2.32 \times L_{\text{flex}}$$

The above relationship is plotted in Figure S3b for a single shank device (solid line) and for a four shank device (interrupted line), where the slope is four times steeper.

How hot does the headstage actually get? The total power dissipation of a 32-channel Intan IC (RHD2132) acquiring wideband (0.1 Hz–10 kHz) signals at peak sampling rate (30 kSamples/s per channel) is 37.3 mW. The temperature of the headstage is dominated by convective cooling and when a single acute PCB is exposed to the ambient air at 25 °C, Newton's law of cooling yields $T_{\text{intan}} = 36.9$ °C, given 298.4 mW of total power dissipation by 8 ICs, 2.4×10^3 mm^2 of PCB area, and a convective heat transfer coefficient for air of 10.45 W$\cdot$m$^{-2}\cdot$K^{-1}. However, when 4 PCBs are closely stacked together to support a 3-D array, the headstage temperature increases, since the PCBs sandwiched in the middle do not effectively exchange heat with the ambient air. A rough upper bound on the PCB temperature reached under these conditions can be obtained by assuming that 4 times the power is dissipated by the surface area of a single PCB. Newton's law of cooling then yields $T_{\text{intan}} = 72.5$ °C, which would require flex cable length of ∼16 mm to limit the rate of heat transfer to the brain below the admissible guideline. Given the design choice of 20 mm cable length, headstage temperatures up to 83 °C can be tolerated.

In the above computations the acute headstage is exposed to the ambient air, but in chronic experiments the headstage is sometimes enclosed in an implant package, thereby limiting heat exchange with the air. We therefore kept a single headstage running inside a closed styrofoam cup and using the Intan chip's internal temperature sensors measured a steady state temperature of ∼40 °C, which would require ∼1.5 mm of cable length for proper thermal insulation. As in the case above, stacking 4 headstages for a chronic 3-D array would require an order of magnitude longer cables.

In summary, providing a high thermal resistance barrier is essential for limiting the rate of heat transfer to the brain below the guideline for implantable devices. In our modular design, the flexible cable fulfills this role. The magnitude of the thermal resistance is proportional to the flex cable length and the design choice of 20 mm offers appropriate thermal barrier for both stacked PCBs supporting 3-D arrays and chronic packaging.

Surgical Procedures and *In Vivo* Recording Methods

Male mice (C57BL/6-E; Strain Code 475; Charles River Laboratories), ∼3 months old, were anesthetized with 1% isoflurane and surgically implanted under aseptic conditions with a light-weight, stainless steel ring embedded in dental cement, which allowed for mechanically stable head-fixation in the recording apparatus, as previously described [S4]. A stainless steel wire was implanted over the right cerebellum and was used as reference. The skull was leveled and the locations of the exposure (a rectangle spanning 2 mm to 2.8 mm lateral and −2 mm to −3 mm posterior from Bregma) were marked on the skull over the left hemisphere. Following surgery, mice were returned to their home cage, maintained on a 12 hour light/dark cycle, and given access to food and water *ad libitum*. Mice were given at least 48 hours to recover before the day of the experiment.

On the day of the experiment mice were anesthetized with 1% isoflurane, head-fixed in the stereotaxic apparatus (Kopf stereotaxic alignment system), and the skull and dura over the designated coordinates were removed. After recovery the mouse was head-fixed on a spherical treadmill within a virtual reality system sitting on an air table (Fig. S4). Motion of the spherical treadmill drove a virtual reality engine that enabled the mouse to navigate a virtual linear track for water reward. The nanoprobe array was attached to a precision manipulator (Luigs & Neumann Junior 4RE 4-axis manipulator) and was

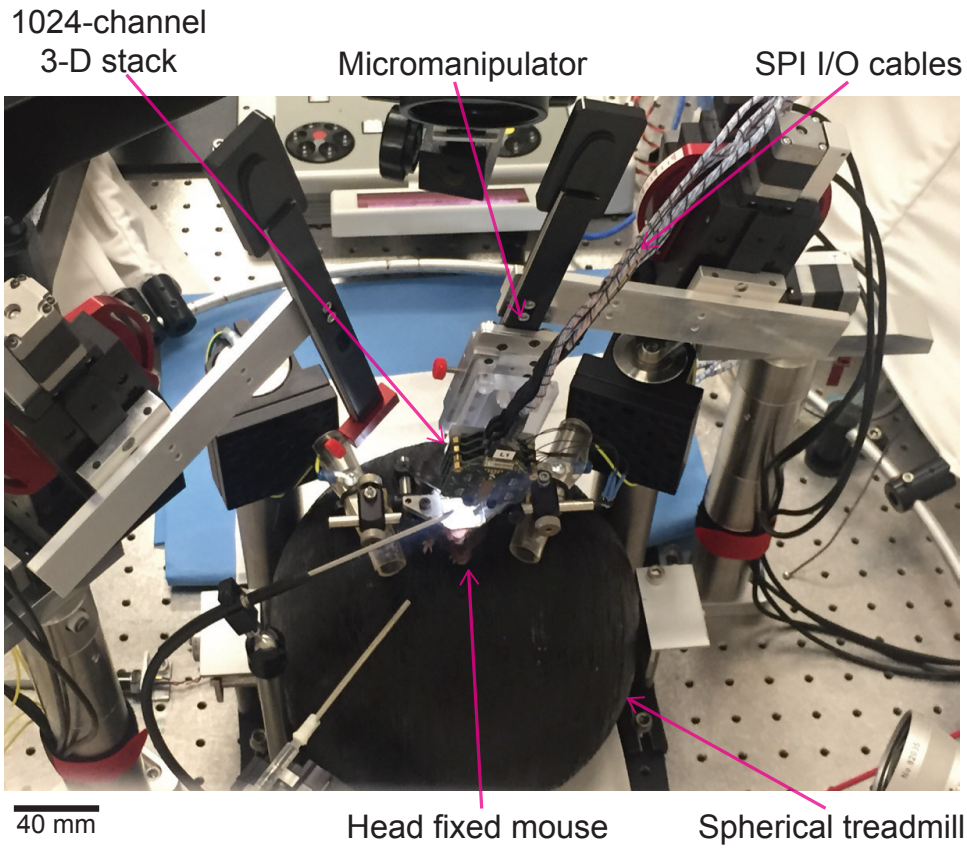

FIG. S4: **Setup for acute *in vivo* recordings.** Head-fixed awake mouse on a spherical treadmill with a 3-D array recording brain activity. The acute headstage stack is seen mounted on the micromanipulator that advances the array to the target area.

gradually lowered (approach angle of 15 degrees from vertical in the coronal plane) until clear electrophysiological signatures of the hippocampal circuit were observed (spiking and ripples in CA1 and CA3, dentate spikes in the dentate gyrus).

All animal procedures were performed in accordance with NIH guidelines and with approval of the Caltech Institutional Animal Care and Use Committee.

Histology

At the end of each experiment the probes were retracted from the brain, the animal was euthanized, and the head was placed in 4% paraformaldehyde (PFA) solution overnight (∼12 hours). The brain was then extracted, kept in 4% PFA for 8 hours, cryo-protected in 30% sucrose for 2 days, and then sectioned into 50-μm-thick slices on a freezing microtome.

Sections were mounted on slides and processed for Hoechst nuclear staining. The tracks left by the shanks were identified across consecutive brain sections to determine the location of each shank in the array (Fig. S5).

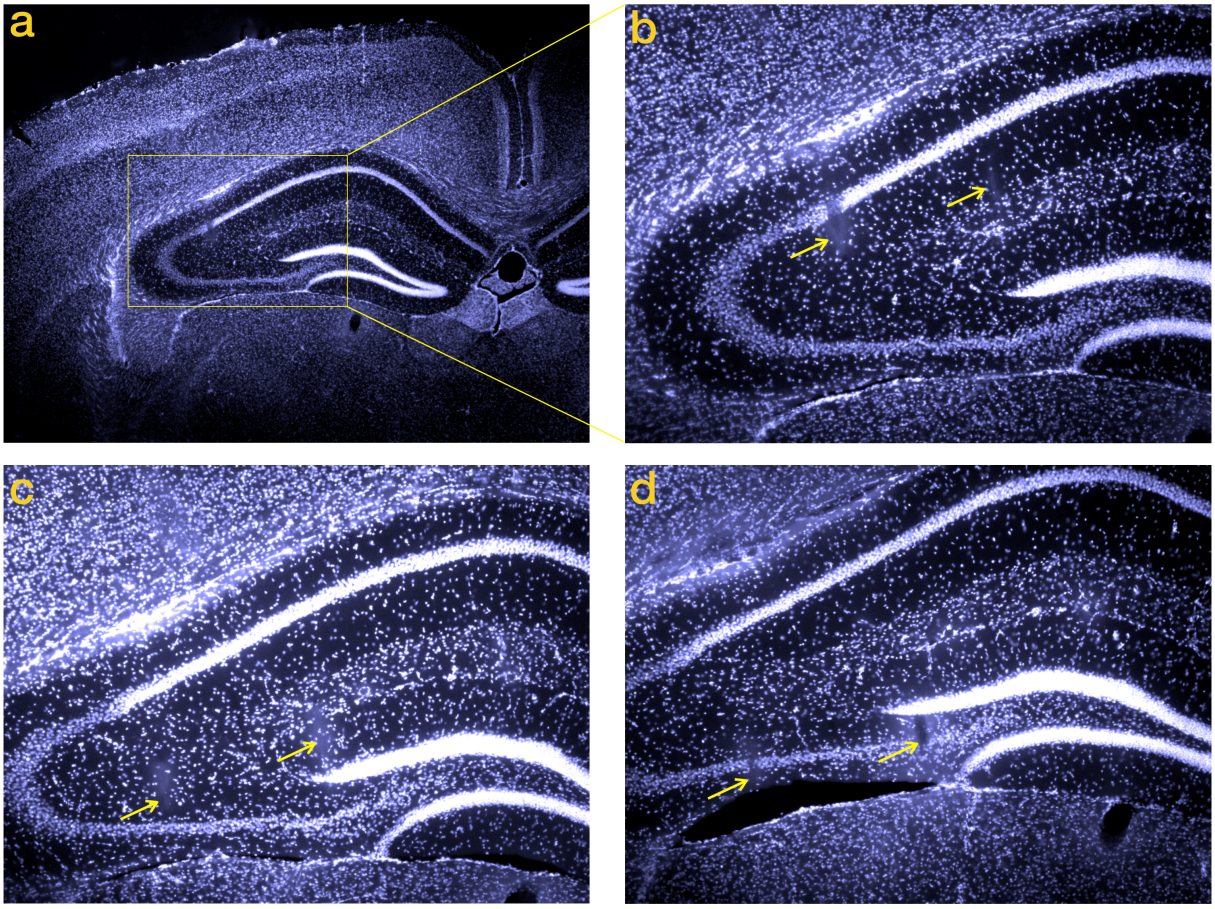

FIG. S5: **Histological identification of electrode locations. (a)** Coronal section (50 μm thick) processed for Hoechst nuclear staining. **(b-d)** Consecutive sections showing identifiable tracks made by two of the four shanks of layer 1 (Figure 5b,c). Section (b) corresponds to the yellow rectangle in (a). The yellow arrows point to the track locations. Notice that the electrode tracks are very thin, given the small dimensions of the shanks, but can be followed across consecutive sections. Identification of a subset of the tracks, in combination with known electrophysiological signatures (spiking and ripples in CA1 and CA3, dentate spikes in the dentate gyrus), enables reliable verification of the location of the 3-D array.

Recording Analysis and Validation

During the *in vivo* evaluation of the recording system, the 3-D array was targeted to the dorsal hippocampus, where the characteristics of the extracellular potential have been thoroughly mapped out. Figures S6 and S7 show data from one 64-electrode shank of the

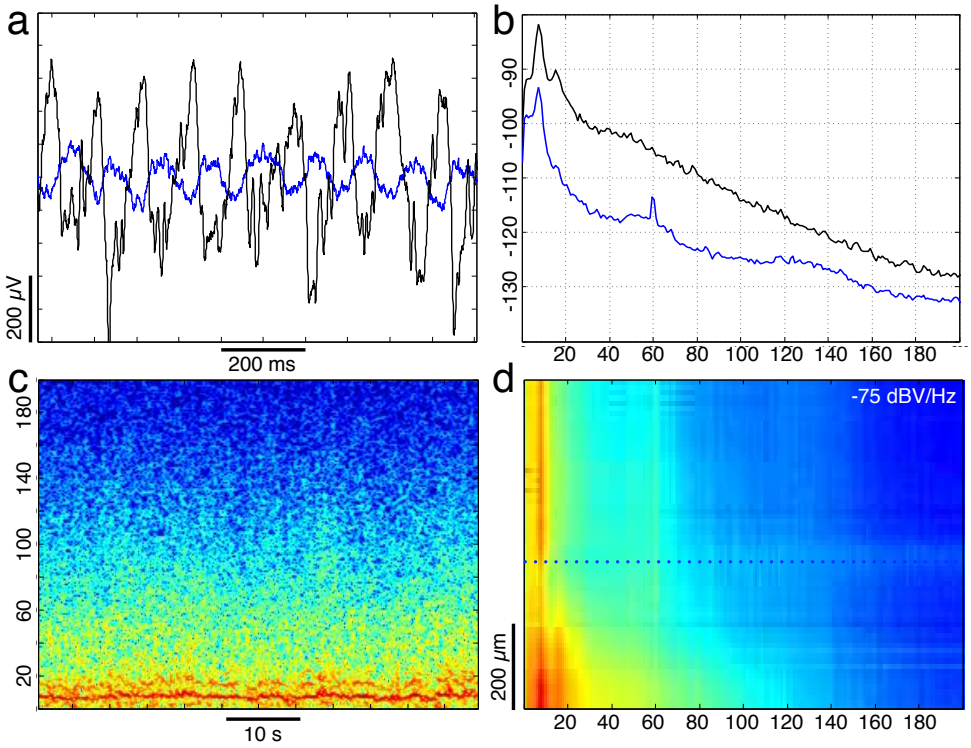

FIG. S6: **LFP spectral content**. **(a)** Local field potentials (LFPs) recorded form the most superficial (blue) and deepest (black) sites on a shank spanning the CA1 pyramidal cell layer of the hippocampus during a 1-second period with prominent theta oscillations. **(b)** Power spectral density (PSD) estimates (in dBV/Hz) of the signals in (a). Notice the prominent theta peaks at ∼9 Hz. **(c)** Spectrogram of the LFP from the deepest site showing sustained theta oscillations over a 1-minute period. **(d)** PSD estimates from each of the 64 sites are vertically stacked (superficial on top) and displayed as a pseudocolor image. Notice prominent theta band throughout with maximal power at the deepest sites (bottom). Dotted line marks the location of the pyramidal cell layer in CA1.

3-D array that approximately straddles the CA1 pyramidal cell layer. During periods of theta oscillations, local field potentials (LFPs) from all sites showed prominent spectral

peaks in the 4-10 Hz theta band, with highest oscillation amplitude observed at the deepest sites near the hippocampal fissure (Fig. S6). Furthermore, the phase of theta oscillations

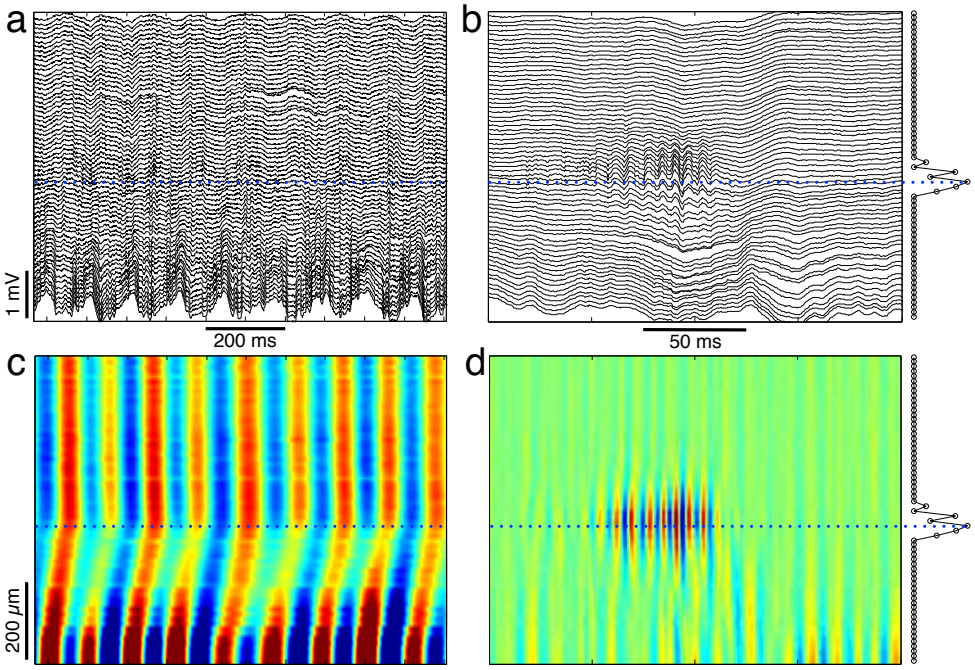

FIG. S7: **Hippocampal LFP features. (a,b)** Local field potentials (LFPs) from 64 sites on a shank spanning the CA1 pyramidal cell layer (blue dotted line) during (a) 1 s period with prominent theta oscillations and (b) 200 ms period containing a ripple. Inset on the right shows the number of high amplitude spikes detected on each site. **(c,d)**. LFPs from the corresponding panels above filtered in the (c) theta or (d) ripple bands and displayed as pseudocolor images. Notice the characteristic shift in the phase of theta oscillations below the CA1 pyramidal cell layer and the limited anatomical extent of ripple oscillations near the cell layer.

was relatively constant above the CA1 pyramidal cell layer and gradually shifted below (Fig. S7a,c), as previously reported [S5, S6]. Finally, ripple oscillations were prominent only near the pyramidal cell layer and were associated with increase in neuronal firing (Fig. S7b,d) [S6, S7]. Thus, LFPs recorded with the nanoprobe 3-D array recapitulate some of the well-known electrophysiological signatures of hippocampal activity.

The recordings obtained with the nanoprobe 3-D array also allow for single unit isolation. This is illustrated in Figure S8, where, for the purposes of comparison, a virtual tetrode is created from four adjacent electrode sites. Notice that the principal components

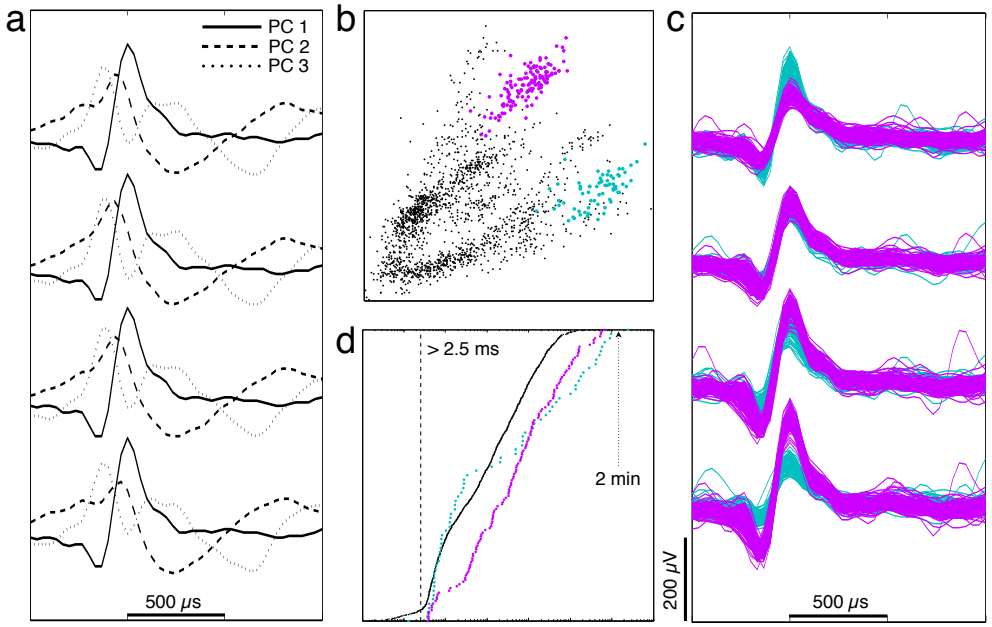

FIG. S8: **Single unit isolation.** **(a)** First three principal components of all candidate spike event waveforms detected on four adjacent sites. **(b)** Scatter plot of PC1 scores from sites 1 and 4. Spikes from two units are seen as separate clusters (cyan and blue). **(c)** Extracellular spike waveforms (inverted polarity) for the two units illustrated in (b). **(d)** Inter spike interval (ISI) empirical cumulant density functions (cdf) for all detected spikes (black) and the two example units.

(PCs) of candidate spike waveforms detected on each of the four sites are very similar and have the characteristic shapes of extracellular action potentials (shown inverted in Fig. S8). Furthermore, the commonly used feature space based on spike waveform PC scores allows for single unit isolation. The spike waveforms and inter-spike interval (ISI) distributions corresponding to two isolated clusters confirm that they likely correspond to single units. About 10 units could be isolated from the virtual tetrode shown in Figure S8, comparable to the performance of a carefully positioned real tetrode.

References

[S1] Maged M Elwassif, Qingjun Kong, Maribel Vazquez, and Marom Bikson. Bio-heat transfer model of deep brain stimulation-induced temperature changes. *J Neural Eng*, 3(4):306–15, Dec 2006.

[S2] C R Davies, F Fukumura, K Fukamachi, K Muramoto, S C Himley, A Massiello, J F Chen, and H Harasaki. Adaptation of tissue to a chronic heat load. *ASAIO J*, 40(3):M514–7, 1994.

[S3] Patrick D. Wolf. Thermal considerations for the design of an implanted cortical brain–machine interface (BMI). In William M. Reichert, editor, *Indwelling Neural Implants: Strategies for Contending with the In Vivo Environment.*, chapter 3. CRC Press/Taylor & Francis, Boca Raton (FL), 2008.

[S4] Brad K Hulse, Laurent C Moreaux, Evgueniy V Lubenov, and Athanassios G Siapas. Membrane potential dynamics of CA1 pyramidal neurons during hippocampal ripples in awake mice. *Neuron*, 89(4):800–13, Feb 2016.

[S5] György Buzsáki. Theta oscillations in the hippocampus. *Neuron*, 33(3):325–40, Jan 2002.

[S6] Evgueniy V Lubenov and Athanassios G Siapas. Hippocampal theta oscillations are travelling waves. *Nature*, 459(7246):534–9, May 2009.

[S7] J Csicsvari, H Hirase, A Czurkó, A Mamiya, and G Buzsáki. Fast network oscillations in the hippocampal CA1 region of the behaving rat. *J Neurosci*, 19(16):RC20, Aug 1999.