Optical magnetic detection of single-neuron action potentials using quantum defects in diamond

SI Appendix

Electrophysiology

Determination of $B^{\mathrm{calc}}(t)$ requires recording the intracellular AP voltage $\Phi_{\mathrm{in}}(t)$. Intracellular microelectrodes are pulled from commercial glass (World Precision Instruments 1B150-4) to a resistance of 20 – 40 M on a pipette puller (Dagan DMZ Universal Puller), filled with 3M KCl, and fitted into a headstage (Axon Instruments HS-2A) connected to an amplifier (Axon Instruments AxoProbe 1B). The microelectrode is mounted to a micro-manipulator for precise insertion into the axon. *M. infundibulum* specimens are dissected as described above but remain pinned in the dissection dish during this measurement so that the microelectrode remains sealed to the axon despite the worm's contractions. AP stimulation occurs as described above. The time trace $\Phi_{\mathrm{in}}^{\mathrm{meas}}(t)$ is digitized (Tektronix TDS2004B), and subsequently low-pass-filtered at 2 kHz. Axon resting potential values Φ_0 of -60 mV to -85 mV and peak AP amplitudes of 100 mV to 110 mV are observed for *M. infundibulum*, consistent with typical values in the literature[1].

For *M. infundibulum*, APs are found to be abolished for values of $f_{\mathrm{stim}} \gtrsim 5$ Hz, and maximal AP amplitudes are realized for $f_{\mathrm{stim}} \lesssim 1$ Hz, both in agreement with previous reports in the literature[2]. For *L. pealeii*, consistent AP stimulation is observed up to $f_{\mathrm{stim}} = 300$ Hz, although eventual axon degradation is observed over ~15 minutes at such high rates. *L. pealeii* exhibits a refractory period following the AP, wherein the potential experiences an overshoot below the resting voltage. The overshoot is not present in *M. infundibulum.*

Electrophysiology for tapered axon studies

For AP stimulation in tapered worm axon studies employing both intracellular and extracellular voltage detection, bipolar platinum microelectrodes (World Precision Instruments PTM23B10) deliver current pulses from an isolated pulse stimulator (A-M Systems Model 2100), to precisely control the location of initiation of APs. For posterior (anterior) stimulation, the bipolar electrode pair is placed within ~2 mm of the end of the worm's tail (head). A schematic of the setup is shown in Fig. S2*C*.

For conduction velocity measurements, two pairs of platinum iridium bipolar recording electrodes (World Precision Instruments PTM123B20KT) are placed 2 mm apart about the mid-point of the worm. One electrode of each pair is placed on the surface of the worm above the axon while the other is displaced laterally 5 mm from the worm to perform a differential measurement. The electrode pairs are attached to a headstage (A-M Systems Model 1800 headstage). The signals are amplified by 100 × by a differential amplifier (A-M Systems Model 1800) and digitized (National Instruments USB-6259) at 100 kHz. The signals are then high-pass filtered at 0.1 Hz, low-pass filtered at 20 kHz, and smoothed in software using a rectangular 7-point moving average. Measurements are taken after ensuring the electrodes are positioned well enough that the coupling to the axon allows a clearly defined peak upon AP stimulation for both channels. The measured conduction velocity is taken as the spatial distance between the pairs of electrodes along the axon (2 mm) divided by the recorded time delay between peaks in the two differential voltage measurements (see Fig. S2*A*, *B*, *D*). We find $v_{c,p}/v_{c,a} = 0.72 \pm 0.11$ (mean $\pm$ s.d. of 1500 pairs of measurements of v_c taken on 15 worms, each with 100 AP trials stimulated from the posterior and 100 from the anterior, see Fig. S2*D*), which corresponds to a p-value of 5.6×10^{-8} relative to the null hypothesis that $v_{c,p} = v_{c,a}$ (student's single sided t-test).

Intracellular AP measurements utilize microelectrodes described above. The intracellular AP voltage $\Phi_{\mathrm{in}}(z_0, t)$ is recorded at a point denoted z_0 at the mid-point of the model axon. Detected signals are digitized (National Instruments USB-6259) at 100 kHz, smoothed using a rectangular 7-point moving average (see Fig. S2*E*), and the temporal derivative is calculated (Fig. S2*F*). We find $\frac{\partial \Phi_p}{\partial t} / \frac{\partial \Phi_a}{\partial t} = 0.97 \pm 0.11$ (mean $\pm$ s.d. of 21 pairs of measurements of $\frac{\partial \Phi}{\partial t}$ taken in 3 worms, each with 7 measured AP trials with posterior stimulation and 7 with anterior stimulation). Combining the results of the intracellular and extracellular measurements, the expected ratio of B_p/B_a is then $v_{c,a}/v_{c,p} \times \frac{\partial \Phi_p}{\partial t} / \frac{\partial \Phi_a}{\partial t} = 1.35 \pm 0.25$.

Tapered axon action potential simulations

Tapered axon AP simulations are performed using the NEURON software[3], a numerical modeling package based on cable theory. The *M. infundibulum* giant axon is modeled as a cylindrically symmetric cable of length $\sim 5\lambda_{\mathrm{AP}}$, (where $\lambda_{\mathrm{AP}} = v_c t_R$ and t_R is the AP impulse rise time) and with a linearly tapered diameter from 400 μm at the anterior to 50 μm at the posterior. The chosen geometric parameters are roughly consistent with axons used in the present magnetic studies[4]. The intracellular conductivity is set to 1.47 $\Omega^{-1}\mathrm{m}^{-1}$, a value typical for *M. infundibulum*[5]. Simulated AP signals are recorded at the mid-point of the model axon so that the sealed-axon boundary conditions can reasonably be neglected. An asymmetry in $\frac{\partial \Phi}{\partial t}$ and v_c is observed depending upon whether the AP is initiated at the posterior (smaller) or anterior (larger) end of the model axon, with $v_{c,p}/v_{c,a} = 0.63$ and $\frac{\partial \Phi_p}{\partial t} / \frac{\partial \Phi_a}{\partial t} = 0.87$. The expected ratio of B_p/B_a from the simulations is then $v_{c,a}/v_{c,p} \times \frac{\partial \Phi_p}{\partial t} / \frac{\partial \Phi_a}{\partial t} = 1.4$.

Transverse sections

Transverse sections shown in Figs. 4*A*, *B* and S1*D* are prepared from a typical *M. infundibulum* specimen, which has been anesthetized and fixed in a solution of 4% paraformaldehyde (PFA) overnight. The fixed worm is mounted to a paraffin block and sectioned, yielding slices of width ~4 μm. The slices are treated with a hematoxylin and eosin stain (H&E stain) to illustrate the tissue structure. Representative slices from three different sections are shown (https://slide-atlas.org/link/mn74xw). By comparison with the literature[4], the observed structure in the sections is verified and the location of the giant axon is confirmed. The sections show a distance from the giant axon center to the skin surface of 900 $\pm$ 200 μm, and a taper in the axon with decreasing diameter from anterior to posterior.

NV-diamond physics

NV color centers are localized quantum defects in diamond consisting of a substitutional nitrogen adjacent to a vacancy in the lattice. The NV center has an $S = 1$ (triplet) ground state with a zero-field splitting of 2.87 GHz between the $m_s = 0$ and $m_s = \pm 1$ spin projections. These states have additional hyperfine structure, which arises from the coupling of the ^{14}N nuclear spin $I = 1$ to the unpaired NV electron spin. A local magnetic field induces Zeeman shifts, lifting the degeneracy of the $m_s = \pm 1$ energy levels. Optically-induced electronic transitions to the excited triplet state and fluorescent decay back to the ground electronic state are mainly spin-conserving[6]. Fluorescent readout and optical polarization of the NV spin state are made possible through a nonradiative decay path from the $m_s = \pm 1$ excited states through metastable singlet states and preferentially to the $m_s = 0$ ground state[6].

Magnetometry method

In contrast to scalar magnetometers, (such as vapor cell magnetometers, proton precession magnetometers, and Overhauser effect magnetometers), which measure magnitude but not direction of a magnetic field, vector magnetometers measure the magnetic field projection onto a fixed axis of the sensing device[7]. Scalar magnetometers have the property that the sensing element tends to align along the magnetic field, and thus the sensor can sense only the magnitude and not the direction. In contrast, SQUIDS, Hall probes, fluxgate magnetometers, and NV-diamond magnetometers are vector magnetometers, since the sensing element is fixed in space and does not align along the direction of the magnetic field. An NV spin is quantized along one of four crystallographic NV symmetry axes in the diamond crystal; and therefore a single NV center senses not the magnitude of the magnetic field but rather the field projection along its own symmetry axis, defined by the line connecting the nitrogen and the vacancy. In a typical NV ensemble with no preferential orientation[8], the four crystallographic NV symmetry axes are equally populated[9]. An ensemble NV-diamond magnetometer can be set up to be sensitive to the magnetic field projection along a single NV axis or along a vector that equally projects onto two or more of the NV axes. The experiments detailed in this work measure the magnetic field projection along the vector that equally projects onto two NV axes; and thus the device operates as a vector magnetometer. Because the four NV axes form a basis set that spans three-dimensional space, a local magnetic field's magnitude and direction can be reconstructed from its measured projections onto the NV axes[10,11].

AP magnetic fields are expected to encircle the axon and be directed perpendicular to the axon axis and hence the direction of AP propagation. In the magnetometer configuration detailed in this work, the sensor records the component B_x of the AP magnetic field parallel to the NV layer surface, and is most sensitive to APs from axons oriented on the diamond surface perpendicular to the x-axis. In this work the axons are oriented roughly linearly on the diamond and perpendicular to two NV axes, maximizing the projection of the AP magnetic field $B(t)$ onto those axes, as shown in Fig. S3*A*, with the projection along the other two NV axes expected to be near-zero.

For an axon oriented at angle α with respect to the direction of maximum sensitivity, the sensed magnetic field is reduced by a factor $\cos(\alpha)$: $B_x(t) = B(t)\cos(\alpha)$. This relationship is demonstrated experimentally in Fig. S4 for different orientations of a current-carrying copper wire with respect to the direction of maximum sensitivity. A single period pulse of a sine wave with frequency 512 Hz is sent through the wire, resulting in a 94 nT peak amplitude magnetic field at the NV sensing layer. For any orientation in the plane except for $\alpha = 90°$, the magnetic field projection along the x-axis is nonzero, with the sign of the measured bipolar magnetic field waveform indicating the projection of the current pulse propagation direction along $\alpha = 0°$. Similarly, when measuring AP magnetic fields, the projection of the AP propagation direction along the x-axis can be determined from the magnetic time trace $B_x(t)$. Moreover, if the expected signal size for $\alpha = 0°$ is known *a priori*, the angle α can be determined from a measurement of $B_x(t)$.

In a next-generation NV-diamond instrument for magnetic imaging of networks of neurons oriented arbitrarily on the diamond surface, the component of $B(t)$ perpendicular to the NV layer at each point on the diamond surface can be sensed, as shown in Fig. S3*B*. The azimuthal AP magnetic field projections would then have opposite sign for measurement points on different sides of the axon, and $B(t)$ would in general have nonzero projection on each of the four NV axes, allowing mapping of AP propagation in the network.

In the present instrument, a modified CW-ESR technique is employed for NV-diamond ODMR, wherein optical NV spin polarization, MW drive, and spin-state readout via LIF occur simultaneously. Continuous green laser excitation at 532 nm polarizes the NV center into the $m_s = 0$ ground state. Applied MWs, when tuned to resonance with the transition between the optically bright $m_s = 0$ spin state and one of the less bright $m_s = +1$ or -1 states, cause NV spin precession into a mixed state and a detectable reduction in LIF. A change in the local magnetic field shifts the ODMR feature and, for near-resonant MW drive, is detected as a change in the fluorescence rate.

A single ODMR feature of Lorentzian lineshape with angular frequency ω_0 (where $\omega \equiv 2\pi f$), linewidth Γ, and contrast $\mathcal{C}$ is detected in fluorescence as $F(\omega) = F_0\left(1 - \mathcal{C}\frac{(\Gamma/2)^2}{(\Gamma/2)^2 + (\omega - \omega_0)^2}\right)$ (see Fig. S5*A*), where F_0 is the fluorescence detected in the absence of MWs, and optical and MW broadening are ignored for simplicity. As the majority of noise in the system has $1/f$ character, greater SNR is achieved by shifting the measurement bandwidth to higher frequency via a lock-in technique, which generates a dispersion-like signal with a characteristic zero-crossing feature: i.e., a rapid change of the lock-in amplifier (LIA) voltage and sign with frequency. The applied MWs are square-wave frequency modulated at f_{mod} (typically 18 kHz) about the center frequency ω_c with frequency deviation ω_{dev}: i.e., $\omega_{\mathrm{MW}}(t) = \omega_c + \omega_{\mathrm{dev}}\mathrm{square}(2\pi f_{\mathrm{mod}} t)$. The collected fluorescence is then $F(\omega_{\mathrm{MW}})$. After demodulation by the LIA with a reference signal $V_{\mathrm{ref}}\sin(2\pi f_{\mathrm{mod}} t)$, the DC output is a dispersion-type signal with a zero-crossing at ω_0:

$$V_{\mathrm{LIA}}(\omega_c, \omega_{\mathrm{dev}}) \propto \frac{F(\omega_c + \omega_{\mathrm{dev}}) - F(\omega_c - \omega_{\mathrm{dev}})}{2} = \frac{V_0 \mathcal{C}}{2}\left(-\frac{(\Gamma/2)^2}{(\Gamma/2)^2 + \left((\omega_c + \omega_{\mathrm{dev}}) - \omega_0\right)^2} + \frac{(\Gamma/2)^2}{(\Gamma/2)^2 + \left((\omega_c - \omega_{\mathrm{dev}}) - \omega_0\right)^2}\right), \quad (1)$$

where V_0 is a prefactor voltage determined by F_0 and the output settings of the LIA. Setting $\omega_{\mathrm{dev}} = \frac{\Gamma}{2\sqrt{3}}$ theoretically maximizes the slope of the zero-crossing $\left.\frac{dV_{\mathrm{LIA}}}{d\omega_c}\right|_{V_{\mathrm{LIA}}=0}$ in the absence of power broadening[12]. Time-varying magnetic fields $B(t)$ are sensed by setting $\omega_c = \omega_0|_{t=0}$ and detecting resonance frequency shifts $\omega_0(t) = \omega_0 + \delta\omega(t)$, where $\delta\omega(t) = \frac{g_e \mu_B}{\hbar} B(t)$, as:

$$V_{\mathrm{LIA}}(t) = V_{\mathrm{LIA}}\left(\omega_0 - \frac{g_e\mu_B}{\hbar}B(t)\right)$$
$$= \frac{V_0\mathcal{C}}{2}\left(-\frac{(\Gamma/2)^2}{(\Gamma/2)^2 + \left(\frac{\Gamma}{2\sqrt{3}} - \frac{g_e\mu_B}{\hbar}B(t)\right)^2}\right.$$
$$\left. + \frac{(\Gamma/2)^2}{(\Gamma/2)^2 + \left(\frac{\Gamma}{2\sqrt{3}} + \frac{g_e\mu_B}{\hbar}B(t)\right)^2}\right)$$
$$\approx -\frac{3\sqrt{3}}{4}\frac{V_0\mathcal{C}}{\Gamma}\frac{g_e\mu_B}{\hbar}B(t).$$

2

The NV spin resonance has three features separated by the hyperfine (HF) splitting of $\Delta\omega_{HF} = 2\pi \times 2.16$ MHz, as shown in Fig. S5*B*. For a single MW frequency sweeping across the features, and again ignoring MW power broadening, we find

$$F(\omega) = F_0\left(1 - \sum_{q=-1}^{1}\mathcal{C}\frac{(\Gamma/2)^2}{(\Gamma/2)^2 + (\omega - (\omega_0 + q\Delta\omega_{\mathrm{HF}}))^2}\right).$$

3

Addressing all three NV HF features simultaneously with three MW frequencies also separated by $\Delta\omega_{\mathrm{HF}}$ yields

$$F(\omega, \Delta\omega_{\mathrm{HF}})$$
$$= F_0\left(1 - \sum_{p=-1}^{1}\sum_{q=-1}^{1}\mathcal{C}\frac{(\Gamma/2)^2}{(\Gamma/2)^2 + ((\omega + p\Delta\omega_{\mathrm{HF}}) - (\omega_0 + q\Delta\omega_{\mathrm{HF}}))^2}\right).$$

4

As displayed in Fig. S5*C*, the observed NV fluorescence signal shows five ODMR peaks. The outer two peaks correspond to one of the three MW frequencies on resonance; the second and fourth peaks correspond to two of the three frequencies tuned to resonance; and the innermost peak corresponds to all three MW frequencies resonantly addressing the HF features. The dispersion signal is then:

$$V_{\mathrm{LIA}}(\omega, \Delta\omega_{\mathrm{HF}}, \omega_{\mathrm{dev}})$$
$$= V_0\left(\sum_{p=-1}^{1}\sum_{q=-1}^{1} -\mathcal{C}\frac{(\Gamma/2)^2}{(\Gamma/2)^2 + ((\omega + p\Delta\omega_{\mathrm{HF}} - \omega_{\mathrm{dev}}) - (\omega_0 + q\Delta\omega_{\mathrm{HF}}))^2}\right.$$
$$\left. + \mathcal{C}\frac{(\Gamma/2)^2}{(\Gamma/2)^2 + ((\omega + p\Delta\omega_{\mathrm{HF}} + \omega_{\mathrm{dev}}) - (\omega_0 + q\Delta\omega_{\mathrm{HF}}))^2}\right).$$

5

In this simple treatment in which MW power broadening is ignored, our measurement technique increases the contrast of the central NV HF feature by a factor of 3. In practice, a contrast improvement factor of ≈ 1.9 is achieved compared to the case of addressing a single HF feature.

The overall measurement contrast is further improved by orienting the bias field B_0 to have equal projection along two NV axes. Projecting along two NV axes doubles the contrast as shown by comparing Fig. S5*C* and *D*, although the angle between the NV axes and $B(t)$ causes the sensitivity improvement to be $2\sin[\theta_{\mathrm{tet}}/2] = 2\sqrt{2/3}$ where $\theta_{\mathrm{tet}} = 109.4712°$ is the tetrahedral bond angle in the diamond lattice.

NV-diamond magnetometer details

The diamond used in this work is an electronic grade (N < 5 ppb) single crystal chip, with rectangular dimensions 4 mm × 4 mm × 500 μm, grown using chemical vapor deposition (CVD) by Element Six. The 13 μm thick top-surface NV sensing layer consists of 99.999% ^{12}C with 27 ppm ^{14}N as evaluated by secondary ion mass spectroscopy, which is irradiated with 4.6 MeV electrons with 1.3 × 10^{14} cm^{-2}s^{-1} flux for 5 hours and subsequently annealed in vacuum at 800 °C for 12 hours. The measured nitrogen-to-NV$^-$ conversion efficiency is ~6%. The diamond is cut so that the 500 μm × 4 mm faces are perpendicular to the [110] crystal axis. The sides are mechanically ground to an optical-quality polish. The diamond is mounted to a 2" diameter, 330 μm thick silicon carbide (SiC) heat spreader via thermal epoxy (Epotek H20E) as shown in Fig. S7*A*. A 2 mm × 25 mm slot in the SiC provides access to the diamond surface for the dissected axon studies.

For NV magnetometry, the diamond sensor is illuminated by 2.75 – 4.5 W of 532 nm laser light (Coherent Verdi V-5) as shown in Fig. S6*B*. Laser light is guided into the diamond via an in-house fabricated UV fused silica coupler, making a ~ 3° angle to the NV layer. A 13 μm thick and 1 mm wide reflective aluminum layer on the diamond surface blocks both excitation light scattered by surface defects and LIF from impinging upon the specimen. A rare earth magnet (1" × 1" × 1" N42 K&J Magnetics) with south pole facing the diamond creates a bias magnetic field B_0 with equal projections of 7 gauss (0.7 mT) along the two NV axes perpendicular to the axon axis, shifting the MW resonance between the $m_s = 0$ and $m_s = 1$ sublevels to ≈ 2.89 GHz.

Fig. S6*A* shows a schematic of the MW setup. A commercial MW source (Agilent E8257D) outputs a single near-resonant frequency, which is square-wave modulated with frequency deviation $\omega_{\mathrm{dev}} = 2\pi \times 360$ kHz at frequency $f_{\mathrm{mod}} = 18$ kHz (Rigol DG1022U). The modulated MWs pass through an isolator (Teledyne Microwave T-2S73T-II) and a -10 dB coupler before mixing via a double balanced mixer (RELCOM M1G) with a 2.16 MHz sinusoid waveform (Stanford Research Systems DS345). The coupled port of the -10 dB coupler is further attenuated by 6 dB and combined (Mini-Circuits ZX-10-2-42-S+) with the mixer output and then sent through a second -10 dB coupler. The coupled output is sent to a spectrum analyzer (Agilent E4405B) while the transmitted output is amplified (Mini-Circuits ZHL-16W-43+), passed through another isolator (Teledyne Microwave T-2S73T-II), a circulator (Pasternack, PE 8401), and a high-pass filter (Mini-Circuits VHF-1200), before delivery to a square 5 mm × 5 mm loop located ≈ 2 mm above the diamond sensor. Slow variations in the NV ODMR resonances, e.g., due to diamond temperature drift, are compensated with ≈ 0.4 Hz feedback to the MW frequency f_{MW}. The ODMR features can also be used to continuously monitor the diamond substrate temperature in real time.

Typically 17 mW (and up to 28 mW) of LIF from the NV-diamond is collected by a 1.4 numerical aperture (NA) aspheric aplanatic oil condenser (Olympus), passed through a 633 nm long-pass filter (Semrock LP02-633RU-25), and imaged onto a biased photodiode (Thorlabs DET100A). The photodiode (PD) is powered by a 12 V lithium ion rechargable battery and is terminated into the RF + DC port of a bias tee (Universial Microwave Component Corporation BT-1000-LS) with bandwidth 10 kHz – 1 GHz. The bias tee DC port is terminated by 50 Ω during experiments; whereas during optical alignment the port is monitored on an oscilloscope to optimize LIF collection. The RF output of the bias tee is amplified by a low noise amplifier (RF Bay LNA-545) and then sent into a LIA (Stanford Research Systems SR850). The LIA gain setting is 200 mV, and the nominal time constant is 30 μs with a 24 dB/octave roll-off, yielding a measured 3 dB cutoff frequency of $f_c = 3.6$ kHz and a

measured equivalent noise bandwidth (ENBW) of $f_{\mathrm{ENBW}} = 4.0$ kHz. The LIA voltage output is expanded by 5× using the LIA expand function, digitized (National Instruments USB-6259) at 250 kHz, and then subsequently divided by 5. The temporary LIA signal expansion is employed to mitigate digitization noise. A ~ 1 nT magnetic field corresponds to a fractional change in the NV LIF of $\frac{\Delta F}{F} \sim 10^{-6}$.

To mitigate correlated magnetic noise from the lab environment, the following procedure is adopted. A lock-in amplifier (Stanford Research Systems, SR830) generates a 60 Hz transistor-transistor logic (TTL) signal phase-locked to the 60 Hz, 120 Volt facility supply power. Each pair of AP stimulations consists of one AP stimulation triggered after a fixed delay T_{delay} relative to a TTL rising edge (following the nominally desired trigger time) and another AP stimulation triggered after the same fixed delay T_{delay} relative to a TTL falling edge (following the nominally desired trigger time). The value of T_{delay} for each pair of AP stimulations is randomly chosen without replacement from an even distribution from 0 to 1/60 seconds. Using this procedure, the magnetic noise is monitored and found to produce no discernable pattern after averaging for 75,000 stimulations under experimental conditions (except with no specimen), as shown in Fig. S9*A*. Allan variance measurements of $B(t)$ taken up to a total measurement time of 100 seconds (with all $B(t)$ filtering off) are consistent with uncorrelated noise. Spurious magnetic signal from the current stimulation is typically well-separated in time from the AP magnetic signal, as shown in Figs. 2*D*, S10*B*, and S11*B*. Furthermore, as discussed below, the magnetometer operates within a factor of two of the predicted sensitivity limited by photon shot noise, which is itself an uncorrelated noise source.

To suppress laser intensity noise near f_{mod}, the 532 nm laser light is sampled and focused on a separate, reference PD (see Fig. S6*B*, *C*). This PD and all electronics (bias-tee, low-noise amplifier, LIA, input into data acquisition system) exactly duplicate the setup of the signal PD and accompanying electronics. The phase of the reference LIA is aligned with the phase of the signal LIA. We find that subtraction (rather than division) of the correlated noise is sufficient to reach the photon shot noise sensitivity limit in the absence of the MWs, in agreement with ref. 13. The detected signal is digitally filtered with an 80 Hz FFT high-pass filter, and with 1-Hz-wide notch stop filters at all 60 Hz harmonics through 660 Hz and at 30 other frequencies above 2 kHz. The experiment achieves sensitivity ~50% above the photon-shot-noise limit, which is discussed in the sensitivity section below.

For the intact worm studies, several changes are made to the experimental apparatus (see Fig. S7*B*). An upgraded aluminum mount (larger than the mount for excised axons) is used to fit the large intact specimens (see Fig. 3*A*). A SiC wafer with no slot is used as a heat spreader. The NV-diamond sensor is therefore offset from the worm exterior by a spacer of thickness 330 μm. MWs are delivered to a 25 μm thick copper foil layer directly on top of the diamond. The Olympus oil aspheric condensor is exchanged for a 0.79 NA air aspheric condensor (Thorlabs ACL25416U-B). Stained transverse sections in Figs. 4*A*, *B* and S1*D* show a typical tissue thickness of ~900 μm from the center of the axon to the worm exterior, consistent with the literature[4,14], although this distance is also noted[4] to be highly variable among different specimens and along a single organism's length. The overall typical distance from the axon center to the diamond sensor is ~1.2 mm, consistent with the measured roughly four-fold magnetic signal reduction compared to excised worm axons, where the distance from axon center to NV detector layer is typically ~300 μm.

Excitation-laser-induced heating of the diamond is measured via NV ODMR frequency shifts to be 2.4 °C/Watt. For the data shown in Fig. 2*C*, *D* (worm B axon and squid A axon) and S10*A*, *B* (worm C axon and squid B axon), the diamond temperature is 21 ± 3 °C. As the excised axons are placed directly against the diamond and thin aluminum layer, we estimate the temperature of both the worm and squid axons to be ~ 21 °C. The live intact organisms (worm D of Fig. 3*B*, worm E of Fig. S10*C*, and worms F, G, and H of Fig. 4*E*, *F*) are separated from the diamond by the SiC heat spreader and are thus at ~ 10 °C during sensing.

Magnetometer calibration

The measured magnetic field $B^{\mathrm{meas}}(t)$ is determined from the output voltage of the LIA, denoted $V_{\mathrm{LIA}}(t)$, by the relation $B^{\mathrm{meas}}(t) = C_{\mathrm{LIA}} V_{\mathrm{LIA}}(t)$, where C_{LIA} is a voltage-to-magnetic-field conversion factor given by

$$C_{\mathrm{LIA}} = \frac{h}{\left.\frac{dV_{\mathrm{LIA}}}{df}\right|_{V_{\mathrm{LIA}}=0} g_e \, \mu_B \sin\left[\frac{\theta_{\mathrm{tet}}}{2}\right]}. \quad (6)$$

Here $\left.\frac{dV_{\mathrm{LIA}}}{df}\right|_{V_{\mathrm{LIA}}=0}$ is the slope of the zero-crossing in V/Hz, g_e is the electron g-factor of the NV$^-$ ground state, $\sin[\theta_{\mathrm{tet}}/2] = \sqrt{2/3}$, and μ_B is the Bohr magneton.

Calibration of the NV-diamond magnetometer is independently verified by applying a known test magnetic field $B_{\mathrm{test}}(t) = B_{\mathrm{test}} \, \mathrm{square}[2\pi f_{\mathrm{test}} t]$ with square wave amplitude B_{test} and frequency f_{test}, and confirming the magnetometer records the correct value for $B^{\mathrm{meas}}(t)$. The test magnetic field is produced by a multi-turn circular current loop (coil) with $N_{\mathrm{turns}} = 7$ and radius $r_{\mathrm{coil}} = 0.0235$ m, located at a distance $z_{\mathrm{coil}} = 0.103$ m from the diamond chip center. The coil is connected in series with an $R_{\mathrm{series}} = 50\ \Omega$ resistor. The value of B_{test} is calculated using the formula

$$B_{\mathrm{test}} = \frac{\mu_0 N_{\mathrm{turns}} I_{\mathrm{coil}} r_{\mathrm{coil}}^2}{2\left[z_{\mathrm{coil}}^2 + r_{\mathrm{coil}}^2\right]^{3/2}}, \quad (7)$$

where I_{coil} is the current in the coil generated by driving a voltage $V_{\mathrm{coil}}(t)$ through the circuit. A 44 mV amplitude square wave yields $B_{\mathrm{test}} = 1.8$ nT, with RMS magnetic field $B_{\mathrm{test}}^{\mathrm{rms}} = B_{\mathrm{test}}$. When this value of B_{test} is applied at frequency $f_{\mathrm{test}} = 110$ Hz, the measured value of $B(t)$ is consistent with the value of B_{test} to better than 5% as shown in Fig. S8*A*. A calibration without harmonics is also performed by applying a 62 mV amplitude sine wave yielding a consistent value of $B_{\mathrm{test}}^{\mathrm{rms}} = B_{\mathrm{test}}/\sqrt{2} = 1.8$ nT.

Magnetic field sensitivity

A magnetometer's sensitivity is defined as $\eta = \delta B \sqrt{T}$, where δB is the magnetic field signal that is as large as the noise, i.e., at SNR=1, after measurement time T[15]. The sensitivity of our NV-diamond magnetometer is evaluated using three methods. In method 1, a test magnetic field $B_{\mathrm{test}}(t) = B_{\mathrm{test}} \sin[2\pi f_{\mathrm{test}} t]$ is applied for $N_{\mathrm{trials}} = 150$, each of time $T_{\mathrm{trial}} = 1$ s, and the measured magnetic field $B^{\mathrm{meas}}(t)$ is recorded. For each trial i the quantity

$$x_i = \frac{1}{T_{\mathrm{trial}}} \int_0^{T_{\mathrm{trial}}} B^{\mathrm{meas}}(t) B_{\mathrm{test}}(t) \, dt \quad (8)$$

is computed. The method 1 sensitivity η_1 is

$$\eta_1 = \frac{B_{\text{test}}^{\text{rms}}\sqrt{2}}{\mu}\sqrt{\frac{1}{N_{\text{trials}}}\sum_{i=0}^{N_{\text{trials}}}(x_i-\mu)^2}\times\sqrt{T_{\text{trial}}}\,,$$

9

where $\mu \equiv \frac{1}{N_{\text{trials}}}\sum_{i=0}^{N_{\text{trials}}} x_i$, the factor of $\sqrt{2}$ accounts for quadrature noise, $B_{\text{test}}^{\text{rms}} = B_{\text{test}}/\sqrt{2}$, and typically f_{test}= 250 Hz. In method 2, $B_{\text{test}}(t)$ is applied for $N_{\text{trials}} = 150$, each of time $T_{\text{trial}} = 1$ s, and $B^{\text{meas}}(t)$ is recorded. The single-sided root-mean-squared (RMS) spectral frequency profile of $B^{\text{meas}}(t)$ is defined to be $\tilde{B}^{\text{meas}}(f)$. The method 2 sensitivity η_2 is

$$\eta_2 = B_{\text{test}}^{\text{rms}}\left\langle\frac{\frac{1}{f_{\text{stop}}-f_{\text{start}}}\int_{f_{\text{start}}}^{f_{\text{stop}}}\tilde{B}^{\text{meas}}(f)df}{\frac{1}{\Delta f}\int_{f_{\text{test}}-\Delta f/2}^{f_{\text{test}}+\Delta f/2}\tilde{B}^{\text{meas}}(f)df}\right\rangle_{N_{\text{trials}}}\times\sqrt{T_{\text{trial}}}\,,$$

10

where $\Delta f = 1/T_{\text{trial}}$, the expected value is taken over N_{trials}, and typically $f_{\text{start}} = 300$ Hz, f_{stop}= 600 Hz, and f_{test}= 250 Hz. In method 3, no test magnetic field is applied and $B^{\text{meas}}(t)$ is recorded for $N_{\text{trials}} = 150$, each of time $T_{\text{trial}} = 1$ s; an example trace is shown in Fig. S8*B*. The sensitivity is then calculated as

$$\eta_3 = \sqrt{\frac{1}{T_{\text{trial}}}\int_0^{T_{\text{trial}}}[B^{\text{meas}}(t)]^2\,dt}\times\frac{1}{\sqrt{2f_{\text{ENBW}}}},$$

11

with $f_{\text{ENBW}} = 4.0$ kHz. In all evaluations of the instrument's magnetic field sensitivity, $\eta_1 \sim \eta_2 \sim \eta_3$ is found, although η_1 converges most slowly and is therefore of limited use. Over 150 trials, η_3 ranges from 15.0 to 15.8 pT/$\sqrt{\text{Hz}}$, while η_2 is 15 ± 1 pT/$\sqrt{\text{Hz}}$. The two values are consistent. We thus conclude the NV-diamond magnetometer sensitivity is 15 ± 1 pT/$\sqrt{\text{Hz}}$, also consistent with a noise floor measurement of $\left|\tilde{B}^{\text{meas}}(2\pi f)\right|$ for $T_{\text{trial}} = 1$ s averaged over $N_{\text{trials}} = 150$, as shown in Fig. S8*C*, *D*.

This realized magnetic field sensitivity agrees with the expected sensitivity for our NV CW-ESR technique limited by photon shot noise and added MW and amplifier noise, as estimated herein. In the limit of low contrast $\mathcal{C}$ of the ODMR feature, defined for CW-ESR as the fractional dip on resonance of collected LIF[16], the photon-shot-noise-limited sensitivity for CW-ESR magnetometry using NV-centers is given by[16]

$$\eta_{\text{ESR}} = \frac{4}{3\sqrt{3}}\frac{h}{g_e\mu_B}\frac{\Delta f}{\mathcal{C}\sqrt{\mathcal{R}}},$$

12

where $\mathcal{R}$ is the photon detection rate (away from resonance, corrected for detector quantum efficiency), Δf is the power-broadened full-width-half-maximum (FWHM) resonance linewidth, and the factor $\frac{4}{3\sqrt{3}}$ comes from a Lorentzian feature's steepest slope. The CW-ESR method employed with the present NV-diamond magnetometer detects along two NV axes as described above, doubling the contrast while reducing the magnetic field sensitivity by the angle factor $\sin[\theta_{\text{tet}}/2] = \sqrt{2/3}$. The shot-noise-limited sensitivity of our magnetometer is given nominally by

$$\eta_{\text{ESR}}^{\text{shot}} = \frac{4}{3\sqrt{3}}\frac{h}{g_e\mu_B}\frac{\Delta f}{\mathcal{C}_2\sin\left[\frac{\theta_{\text{tet}}}{2}\right]\sqrt{\mathcal{R}}},$$

13

where $\Delta f = 1.1 \pm 0.1$ MHz is the measured linewidth; and $\mathcal{C}_2 = 5.3 \pm 0.1\%$ is the ESR contrast when sensing along two NV axes, which is measured in the absence of modulation while addressing all three hyperfine features. The detected photon rate $\mathcal{R}$ is defined in terms of the photoelectron current $q\mathcal{R} = V_{\text{sig}}/R_L$, where q is the elementary charge and $V_{\text{sig}} = 400$ mV is the typical signal PD voltage after $R_L = 50\,\Omega$ termination. This idealized shot-noise-limited CW-ESR sensitivity is found to be 3.1 pT/$\sqrt{\text{Hz}}$.

In practice, several factors diminish the sensitivity: first, the reference PD adds in quadrature an equivalent amount of shot noise, increasing the sensitivity by a factor $\mathcal{P}_{\text{ref}} = \sqrt{2}$; second, the slope is reduced with respect to the steepest slope of a Lorentzian due to the other nearby power-broadened hyperfine features, resulting in a sensitivity cost of $\mathcal{P}_{\text{slope}} = 1.19$. Taking these factors into account yields a shot–noise-limited CW-ESR sensitivity of 5.2 pT/$\sqrt{\text{Hz}}$.

Furthermore, in our square-wave modulated CW-ESR implementation, the contrast is reduced by an empirical factor $\mathcal{P}_{\text{mod}} \approx 1.6$, as shown in Fig. S8*F*, due to the finite cycling time of the NV center quantum states[17,18] and the loss of signal in higher harmonics resulting from demodulation with a sinusoidal lock-in frequency waveform[12]. The LNA-545 amplifier's noise figure of 1.8 increases the noise level by $\mathcal{P}_{\text{ampl}} \approx 1.23$. Excess noise at the 18 kHz modulation frequency increases the noise by $\mathcal{P}_{\text{excess}} \approx 1.16$, while application of MWs further increases the measured noise level by $\mathcal{P}_{\text{MW}} \approx 1.16$, as shown in Fig. S8*E*. These factors raise (i.e. worsen) the expected magnetic field sensitivity to $\eta \approx \mathcal{P}_{\text{MW}}\mathcal{P}_{\text{excess}}\mathcal{P}_{\text{ampl}}\mathcal{P}_{\text{mod}}\mathcal{P}_{\text{slope}}\mathcal{P}_{\text{ref}}\eta_{\text{shot}} \approx 13.8$ pT/$\sqrt{\text{Hz}}$, which agrees to within 8% of the measured 15 ± 1 pT/$\sqrt{\text{Hz}}$ for the data shown in Fig. S8*B-E*.

To confirm magnetometer sensitivity near the photon shot noise limit in the absence of applied MWs, we measure the RMS noise in V_{LIA} for a range of power incident on the PD, at both 18 kHz and 90 kHz modulation frequencies, as shown in Fig. S8*E*. Data are fit to the function $y = \sqrt{a + bx + cx^2}$. For 90 kHz modulation, the fit parameters are $a = (7.9 \pm 0.1) \times 10^{-2}$, $b = (2.2 \pm 0.02) \times 10^{-2}$, and $c = (-3.0 \pm 0.8) \times 10^{-6}$; for 18 kHz modulation the fit parameters are $a = (8.3 \pm 0.2) \times 10^{-1}$, $b = (2.4 \pm 0.1) \times 10^{-2}$, and $c = (0.5 \pm 3.6) \times 10^{-6}$. The fits for both cases suggest the measured noise is dominated by the bx term, as expected for a shot-noise-limited measurement. The measured noise agrees with expected photoelectron shot noise from signal and reference channels plus LNA-545 amplifier noise for equivalent noise bandwidth $f_{\text{ENBW}} = 4.0$ kHz.

The fundamental sensitivity limit for spin-based magnetometers is given by the noise intrinsic to quantum projection. For a sample of N electronic spins with characteristic dephasing time T_2^*, the spin-projection-noise-limited sensitivity is[7]:

$$\eta_q = \frac{\hbar}{g_e\mu_B}\frac{1}{\sqrt{NT_2^*}}.$$

14

The sample used in this work has a total NV density $\sim 3 \times 10^{17}\,\text{cm}^{-3}$ and no preferential orientation[8]. The density of NVs used to sense AP magnetic fields is reduced by a factor of two, as the AP magnetic field projects along only two NV axes. The illumination volume is $\sim 13\,\mu\text{m} \times 200\,\mu\text{m} \times 2\,\text{mm} \approx 5 \times 10^{-6}\,\text{cm}^3$, so the number of probed NV spins is $N \sim 8 \times 10^{11}$ with $T_2^* \approx 450$ ns. Using these values along with the NV electron's gyromagnetic ratio $\gamma = g_e\mu_B/\hbar = 1.761 \times 10^{11}\,\text{s}^{-1}\text{T}^{-1}$ gives a spin projection noise estimate for our sample volume of ~10 fT/$\sqrt{\text{Hz}}$. At ~1500 times better than the present nearly photon-shot-noise-limited sensitivity, there is much promise for significant gains in magnetometer sensitivity through use of pulsed magnetometry, optimized NV-

diamond samples, and quantum-assisted techniques, as discussed below.

Temporal resolution

Temporal resolution of the NV-diamond magnetometer is tested by applying a test magnetic field $B_{\text{test}}(t) = B_{\text{test}}\,\text{square}[2\pi f_{\text{test}} t]$ with $B_{\text{test}} \approx 57$ nT and $f_{\text{test}} = 1$ kHz, and measuring the 10% – 90% rise time of $B^{\text{meas}}(t)$, denoted by $\tau_{10/90}$. Using $f_{\text{mod}} = 60$ kHz, $\tau_{\text{LIA}} = 10$ µs, and 6 dB/octave roll-off (yielding a measured $f_{\text{ENBW}} = 33$ kHz), $\tau_{10/90} = 32\ \mu$s is observed as shown in Fig. S9*B*, which displays both real-time and averaged $B^{\text{meas}}(t)$ traces that are FFT low-pass filtered at 45 kHz. All AP data presented in this paper are acquired using $f_{\text{mod}} = 18$ kHz, $\tau_{\text{LIA}} = 30$ µs, and 24 dB/octave roll-off, which gives $\tau_{10/90} \sim 100$ µs. Note that higher values of f_{mod} reduce NV spin-state contrast, an effect previously observed in refs. 15, 17, 18 and shown here in Fig. S8*F*. When operating with a temporal resolution higher than 40 µs, the magnetic field sensitivity of the present instrument is reduced by a factor of ~1.6 with respect to standard running conditions. With pulsed Ramsey-type schemes[19], to be employed in a next-generation NV-diamond magnetic imaging system, the time resolution is expected to be significantly enhanced without loss of NV contrast. The temporal resolution of a Ramsey scheme is in practice limited by the repetition rate of the pulse sequence. For typical initialization and readout times of 1 µs and 300 ns respectively, and for a 450 ns T_2^*-limited spin precession time, we anticipate a temporal resolution of ~2 µs. Temporal resolution can be further increased at the expense of sensitivity by reducing the initialization and spin precession times, down to a theoretical limit given by the ~200 ns NV singlet state lifetime[20].

Action potential signal-to-noise ratio

The SNR of an AP magnetic field data set is calculated using (i) the peak-to-peak detected AP signal from an averaged set of N_{avg} measurements and (ii) the standard deviation of the time trace in a section of the same data set in which no AP is present. The single-shot SNR is calculated by dividing the SNR of the averaged data by $\sqrt{N_{\text{avg}}}$. For excised axon studies, $t = 0$ corresponds to the beginning of the stimulation pulse. For averaging data in intact organism studies, traces are aligned in time using a digital trigger set on either the maximum or minimum of the extracellular AP voltage signal $\Phi_{\text{ex}}(t)$; this alignment compensates for specimen contractions resulting in variable propagation delays and thus prevents smearing out of the averaged signal. The SNR is further quantified for the specimens studied in Fig. 2*C*, *D* (worm B axon and squid A axon) with a series of technical replicates, which are shown in Fig. S11*A*, *B*.

To maximally improve the SNR of a known expected signal in the presence of white noise, a matched filter can be shown to be the optimal linear filter[21]. For a detected signal $x(t)$ containing an expected signal and additive noise, the matched filtered signal $y(t)$ is given by the convolution

$$y(t) = \int_0^t h(t - t')x(t')dt',$$

15

where $h(t)$ is the time-reversed trace of the expected signal. The matched filter is constructed from the data shown in Fig. 2*C* with $N_{\text{avg}} = 600$. This trace is high-pass FFT filtered at 80 Hz to prevent non-DC values due to drift from being interpreted as signal. The trace is then zeroed for all times except a 1.4 ms window that includes the full detected AP signal, time-reversed, and then taken as the expected signal $h(t)$ for the matched filter. This filter is applied to the four consecutive sets of 150 averages contained in the larger data set (shown in Fig. S11*A*). The SNR of each of these filtered traces is improved to be between 14.5 and 16, indicating that the SNR of a single AP event after filtering is 1.2 ± 0.1 (mean ± s.d. for four magnetic measurements, each with $N_{\text{avg}} = 150$).

Systematic checks

As summarized in Table S1, we present the results of multiple tests performed to verify that the observed $B^{\text{meas}}(t)$ arises solely from an axon AP (i.e., intracellular axial current): (i) observation of a non-zero $B^{\text{meas}}(t)$ signal requires successful AP stimulation and propagation as determined by electrophysiology measurements of the extracellular action potential $\Phi_{\text{ex}}(z, \rho, t)$, and the observed $B^{\text{meas}}(t)$ is synchronized with $\Phi_{\text{ex}}(z, \rho, t)$ to within experimental error; (ii) crosstalk ('pickup artifacts') during data acquistion between the recorded $\Phi_{\text{ex}}(z, \rho, t)$ and $V_{\text{LIA}}(t)$ is ruled out through varying the recording electrode placement and observing no change in $B^{\text{meas}}(t)$; (iii) the origin of the NV-observed $B^{\text{meas}}(t)$ signal is demonstrated to be magnetic by switching to an LIA voltage zero-crossing with slope $\frac{dV_{\text{LIA}}}{df}$ of opposite sign (see Fig. S5), and observing inversion of $B^{\text{meas}}(t)$; (iv) similarly, inverting the phase of the LIA reference signal ϕ_{LIA} by 180° produces the same result, also confirming the magnetic origin of the signal sensed by the NV ensemble; and (v) time-varying magnetic fields from motional artifacts, e.g., specimen-induced instrument motion in the presence of a gradient in the bias field B_0, are ruled out by reversing the orientation of the permanent magnet and observing inversion of $B^{\text{meas}}(t)$.

An investigation of whether muscle presence affects the observed $B^{\text{meas}}(t)$ in intact worm studies is described herein. As shown in Fig. S12*A*, specimens sometimes display an extended duration (~5 – 10 ms typical), temporally delayed (2 – 20 ms typical) extracellular voltage signal, which is eliminated upon perfusion of 10% ethanol, suggesting a muscular signal origin. The delayed signal is easily resolved from the neuron extracellular AP signal, which exhibits shorter delay after stimulation and ~ 1 ms duration. Fig. S12*B* shows extracellular voltage measurements from an intact worm after 15, 30, 45 and 60 repeated stimulations at 0.5 Hz. The delayed extracellular signal is observed to decrease with the number of consecutive stimulations, an effect attributed to muscle fatigue and consistent with independent observations in the literature[22]. Additionally, the shape of $B^{\text{meas}}(t)$ is observed to be consistent between excised axons (Figs. 2*C*, S10*A*) and intact worms (Figs. 3*B*, 4*E*, *F*, and S10*C*), with both found to be in qualitative agreement with predicted $B^{\text{meas}}(t)$ (Fig. 2*B*) derived from intracellular voltage measurements on excised axons. As discussed previously, the amplitude decrease of $B^{\text{meas}}(t)$ in intact worms versus excised axons is attributed to greater standoff distance between sensor and axon location in intact worms. Thus, due to the ease of fatigue, the temporal delay, and the qualitative and temporal agreement of expected and observed shapes of $B^{\text{meas}}(t)$, the presence of the muscle is determined to be unlikely to affect the observed traces of $B^{\text{meas}}(t)$ for intact worms.

Extended duration action potential sensing

For the long-term sensing data shown in Fig. S10*C*, the specimen (intact worm E) is prepared and clamped to the apparatus as described above for intact organism studies. The worm is magnetically monitored for > 24 hours in the presence of applied MWs and laser illumination of the diamond. Following this duration, the magnetic AP signal $B^{\text{meas}}(t)$ is measured to have an amplitude consistent with AP signals of specimens studied over shorter durations (intact worms D, F, G, and H). Physical stimulus applied to the specimen after >24 hours further confirms its responsiveness

and health.

Effects of microwave fields on organisms

The observed temperature increase at the diamond sensor with MWs on versus off is 1 °C or less, with no observable damage to the biological specimens[23]. In particular, no adverse effect of the MWs is observed on the long-term health of intact worm E during the >24-hour study.

Effects of static magnetic fields on organisms

Theoretical calculations predict a static bias field of 24 T is needed to noticeably alter neural conduction[24]. FDA guidelines allow magnetic fields up to 8 T for humans subjected to clinical magnetic resonance imaging[25]. Thus neither the physical magnet itself nor the modest magnetic field of 7 G (0.7 mT) at the location of the diamond is expected to result in deleterious effects to the specimens studied in this or future work.

Expected magnetic field sensitivity in next-generation instrument

In this section we estimate the magnetic field sensitivity of a next-generation magnetic imaging instrument based on existing demonstrated magnetometry techniques and anticipated advances in diamond growth. A next-generation instrument will likely employ pulsed magnetic field sensing schemes, such as Ramsey-type sequences, which do not suffer from laser and MW power broadening of the NV-diamond ODMR features and thus allow for higher contrast and enhanced sensitivity per volume than CW-ESR[19]. A Ramsey scheme with free precession time τ, and optical and MW initialization and readout times t_I and t_R, has a shot-noise-limited sensitivity[9,26,27] of

$$\eta_{\mathrm{Ramsey}} = \frac{\hbar}{g_e \mu_B} \frac{\sqrt{t_I + \tau + t_R}}{\tau} \frac{1}{\mathcal{C}' e^{-(\tau/T_2^*)^p} \sqrt{\beta}},$$

16

where $\beta \approx \mathcal{R} t_R$ is the average number of photons collected per measurement, $\mathcal{C}'$ is the fluorescence contrast at $\tau = 0$, and the exponential factor $e^{-(\tau/T_2^*)^p}$ accounts for contrast degradation due to NV spin dephasing with characteristic time T_2^*. For broadening mechanisms that produce Lorentzian lineshapes, $p = 1$; and for Gaussian lineshapes, $p = 2$. For typical measurements on large ensembles, we may assume $p = 1$[28]. We note that $\mathcal{C}'$ is defined as the difference in collected LIF between the maximum and minimum of a Ramsey fringe divided by the sum of the maximum and minimum collected LIF[9]. This definition differs from the earlier definition of contrast $\mathcal{C}$ used for CW-ESR as $\mathcal{C}' = \frac{\mathcal{C}}{2-\mathcal{C}}$. Depending upon the values of t_I and t_R, the sensitivity is typically optimized for $T_2^*/2 \leq \tau < T_2^*$.

Here we calculate the sensitivity improvement for both the diamond used in the present work and a future anticipated diamond chip with improved sensing parameters. For both estimates we assume a fluorescence contrast per NV axis of $\mathcal{C}' = 3.9\%$, a collection efficiency of CE = 0.2 (corresponding to the collection efficiency of a lossless 1.49 NA oil objective[29] and including a reflective layer on the NV diamond surface to further double the collected LIF), and an excitation intensity near saturation leading to a time-averaged photon emission rate of $\Gamma_{\mathrm{ph}} = 1/(200\ \mathrm{ns})$ per NV. Employing double-quantum magnetometry[30,31] increases the magnetic precession rate by 2 ×, which effectively reduces T_2^* from its native value by 2 ×, so that $T_{2,\mathrm{DQ}}^* = T_2^*/2$. Double-quantum magnetometry can also provide common-mode rejection of noise due to strain and temperature inhomogeneities, promising a further sensitivity enhancement, which is not explicitly accounted for here. Both estimates also presume simultaneous interrogation of all four NV axes, which provides a sensitivity enhancement of $2\sqrt{4/3}$ (= $4\cos[\theta_{\mathrm{tet}}/2]$). Combining these enhancement factors into Eqn. 16 yields

$$\eta_{\mathrm{DQ,4-axis}} = \frac{1}{2} \times \frac{1}{2\sqrt{4/3}} \frac{\hbar}{g_e \mu_B} \frac{\sqrt{t_I + \tau + t_R}}{\tau} \frac{1}{\mathcal{C}' e^{-(\tau/T_{2,\mathrm{DQ}}^*)} \sqrt{\beta}}.$$

17

The diamond used in the present work has native $T_2^* = 450$ ns, [N] = 27 ppm, and ~6% conversion efficiency from [N] to [NV$^-$]. In an illumination volume of 1 μm^3, $n_{\mathrm{NV}} = 3.0 \times 10^5$ NV spins are addressed. For simplicity these estimates assume negligible light losses in the optical system so that the photon collection rate from a 1 μm^3 volume is $\mathcal{R} = \mathrm{CE}\, n_{\mathrm{NV}}\, \Gamma_{\mathrm{ph}} = 3.0 \times 10^{11}$ photons per second. We assume typical values of $t_I = 1$ μs and $t_R = 300$ ns. Using these parameters in Eqn. 17 with $\tau = T_{2,\mathrm{DQ}}^* = T_2^*/2 = 225$ ns, the expected volume-normalized sensitivity of an enhanced sensitivity magnetometer using the present diamond chip is 1.57 nT μm$^{3/2}$Hz$^{-1/2}$. This represents more than 20-fold improvement over the present instrument, with bulk sensitivity 15 pT/$\sqrt{\mathrm{Hz}}$ over a 13 μm × 200 μm × 2000 μm = 5 × 10^6 μm^3 volume, leading to a volume-normalized sensitivity of 34 nT μm$^{3/2}$Hz$^{-1/2}$.

For a next-generation diamond chip, we assume the following parameters: $T_2^* = 21$ μs, $t_I = 2$ μs, $t_R = 300$ ns, [N] = 2 ppm, and 50% conversion efficiency from [N] to [NV$^-$], such that in an illumination volume of 1 μm^3, $n_{\mathrm{NV}} = 1.76 \times 10^5$ and $\mathcal{R} = 1.76 \times 10^{11}$ photons per second. Using these parameter estimates in Eqn. 17 and using $\tau = T_{2,\mathrm{DQ}}^*/2 = T_2^*/4 = 5.25$ μs, the expected volume-normalized sensitivity is 118 pT μm$^{3/2}$Hz$^{-1/2}$, representing a nearly 300× improvement over the present instrument. Further sensitivity enhancements may be possible through quantum-assisted techniques[27,32,33], or even using other solid-state color defects[34].

Simple magnetic model of action potential

The magnetic field produced by an axon AP, denoted $B_{\mathrm{axon}}(z, \rho, t)$, can be derived from the intracellular AP voltage $\Phi(z, \rho, t)$, where z and ρ denote the axial and radial coordinates respectively, using a simple model that agrees with more complex cable theory[35,36]. The axon is modeled as a conducting wire; hence the magnetic field is $B_{\mathrm{wire}} = (\mu_0 I)/(2\pi\rho)$, with axial current I due to the propagating AP. From Ohm's Law, the wire's current density is $J = -\sigma \nabla \Phi(z, \rho, t)$, where σ is the electrical conductivity. For a uniform cylindrical wire of radius r_a, the axial current may be expressed as $I = \pi r_a^2 J_z = -\pi r_a^2 \sigma \frac{\partial \Phi(z,\rho,t)}{\partial z}$. For constant values of conduction velocity v_c, the equality $\frac{\partial \Phi(z,\rho,t)}{\partial t} = -v_c \frac{\partial \Phi(z,\rho,t)}{\partial z}$ holds, where v_c is defined by convention to be positive[35]. Substitution then yields $I = \frac{\pi r_a^2 \sigma}{v_c} \frac{\partial \Phi(z,\rho,t)}{\partial t}$. At distances close to the axon surface where $\rho \sim r_a$, return currents outside the axon are minimal[35], and fringing effects from the finite axon length can be ignored, yielding

$$B_{\mathrm{axon}}(z, \rho, t) = \frac{\mu_0 r_a^2 \sigma}{2 v_c \rho} \frac{\partial \Phi(z, \rho, t)}{\partial t}.$$

18

Defining $s \equiv \frac{\mu_0 r_a^2 \sigma}{2 v_c \rho}$ gives $B_{\mathrm{axon}}(z, \rho, t) = s \frac{\partial \Phi(z,\rho,t)}{\partial t}$, where s depends only on geometric and electrophysiological quantities. For the data shown in Fig. 2*A-C*, good proportionality is found between $B^{\mathrm{meas}}(t)$

and $\frac{\partial \Phi^{\text{meas}}}{\partial t}$ with $s^{\text{meas}} = 7.6 \pm 1$ pT/(V/s) (mean $\pm$ s.d. for four magnetic measurements shown in Fig. S11*A*, each with $N_{\text{avg}} = 150$).

Accurate calculation of s from first principles is nontrivial[37], since r_a, ρ, σ, and v_c have substantial uncertainties. $r_a = 200 \pm 75$ μm is determined from stained transverse sections of *M. infundibulum* (see Fig. 4), with large variations observed in axon size (up to 50%) among otherwise similarly sized specimens, as also noted in ref. 12. For the excised axon studies, only the ventral nerve cord containing the giant axon is isolated, and there is also residual connective tissue around the axon: hence an estimate of $\rho = r_a + 100$ μm $\pm$ 100 μm is used. We take $\sigma = 1.47 \pm 0.5$ Ω^{-1} m^{-1} for *M. infundibulum*[5], given the significant variation (50%) in axoplasm conductivity reported for *L. pealeii*[37]; and $v_c = 9 \pm 4$ m/s (mean $\pm$ s.d. of 105 AP measurements taken on 7 worms, with 15 AP stimulations per worm) based on two-point electrophysiology measurements of v_c (described in the Methods) for similar representative-sized worms under posterior stimulation. With these values for r_a, ρ, σ, and v_c, we extract $s^{\text{calc}} = 13.7 \pm 10$ pT/(V/s), which is in agreement with the experimentally derived value.

Determination of axonal currents from magnetic field and voltage measurements

Because axonal currents effect changes in the membrane potential of passive electric structures[38] understanding and characterizing axonal currents into and out of the soma is expected to increase understanding of computation and summation mechanisms in passive membranes. Thus, measurements of axonal currents can provide insight into electrical dynamics within both active and passive neuronal structures.

Under a simple magnetic model assuming a cylindrical axon with uniform conductivity, the current density J_a is uniform Ampere's law dictates $J_a(z,t) = \frac{2B(r,z,t)}{\mu_0 r}$, where we have implicitly made the common assumption[39] that the magnetic permeability μ in biological tissue is not materially different from the vacuum permeability μ_0. From a measurement of $B(r,z,t)$, the only remaining parameter required to determine the current density is the radial distance to the neuron center *r*.

In contrast, extraction of $J_a(z,t)$ from a voltage measurement requires the knowledge of multiple additional parameters. When Ohm's law is applied to neuronal structures[40] the axonal current can be defined as $J_a(z,t) = \sigma \nabla \Phi(z,t)$, where σ is the internal conductivity and $\Phi(z,t)$ is the intracellular voltage. As discussed in the previous section, for axially symmetric axons the spatial gradient may be related to the temporal derivative by $\frac{\partial \Phi}{\partial z} \approx -\frac{1}{v_c}\frac{\partial \Phi}{\partial t}$ as long as the conduction velocity v_c is well known and constant. Thus, both conduction velocity and electrical conductivity must be well characterized to relate voltage measurements to axonal current. However, without prior knowledge of v_c, a numerical spatial gradient of the intracellular voltage would need to be calculated, which would require multiple simultaneous, spatially separated measurements of the active neuron. Such measurements necessitate high spatial and temporal resolution to extract J_a[41] along with large SNR, as the calculation of spatial gradients of a measured field map amplifies high spatial-frequency noise[42,43,44]. Furthermore, the measured extracellular signal is known to provide an even poorer reconstruction of the intracellular action potential due to variations and uncertainties in bath conductivities and in the effective resistance and capacitance of the measurement probes[37].

Estimate of magnetic signals from mammalian axons

To evaluate the feasibility of NV-diamond magnetic sensing of small mammalian neurons, a crude estimate of the AP magnetic signal size is made for Purkinje neurons using our simple model. We use $\sigma = 0.66\ \Omega^{-1}\text{m}^{-1}$, an average of the values 0.44 Ω^{-1} m^{-1} from ref. 45, 0.87 $\Omega^{-1}\text{m}^{-1}$ from ref. 46, and 0.67 $\Omega^{-1}\text{m}^{-1}$ from ref. 47. We use $\frac{\partial \Phi}{\partial t} = 339$ V/s, an average of 300 V/s from ref. 48, 367 V/s from ref. 49, and 350 V/s from ref. 50. We use $v_c = 0.25$ m/s, an average of 0.24 m/s from ref. 48 and 0.25 m/s from ref. 51. For $r_a = 1$ μm, 2 μm, and 3 μm we calculate a peak magnetic field of $B_{\text{axon}}^{\max}$ = 0.6 nT, 1.2 nT and 1.7 nT respectively at the axon surface. Note that conduction velocity is expected to be correlated with axon radius, which is not accounted for in this calculation. Hence, this calculation is intended only as a rough estimate, as there can be significant variation in the geometric and electrophysiological details of Purkinje neurons[47].

Also note that the average mammalian neuron AP magnetic field sensed over a layer of NV centers of nonnegligible thickness is somewhat reduced. To estimate the field detected by an NV layer in a next-generation magnetic imager, we rely on the axon model[35,36]. The scaling of magnetic field with distance is parameterized by the AP wavelength $\lambda_{\text{AP}} = v_c t_R$. For unmyelinated mammalian axons, $\lambda_{\text{AP}} \sim 1$ mm[52,35]. At sensor-to-source standoff distances much less than λ_{AP}, external return currents do not attenuate the magnetic signal appreciably and axon AP magnetic fields fall off as $1/\rho$[36]. Let the average field at the location of an NV center positioned at the neuron surface directly beneath the neuron (i.e., $\rho = r_a$) be $B(r_a,t) = B_0(t)$. The average field sensed by a layer of NVs of thickness d directly below the neuron surface ranging from $\rho = r_a$ to $\rho = r_a + d$ is then $B_{\text{avg}}(t) = B_0(t)\frac{r_a}{d}\int_{r_a}^{r_a+d}\frac{d\rho}{\rho} = B_0(t)\frac{r_a}{d}\ln\left(\frac{r_a+d}{r_a}\right)$. For example, for an AP signal that reaches a peak value of 1 nT at the surface of an axon, where $r_a = 5$ μm, the average magnetic field detected by an NV layer of thickness $d = 5$ μm below the axon is 0.69 nT.

In addition, the detected signal falls off in the lateral dimensions for NVs not positioned directly under the center of the axon. Taking the positive z-direction to be the propagation direction of the AP, and transforming from cylindrical to Cartesian coordinates ($\rho^2 = x^2 + y^2$, $\tan(\theta) = y/x$), the magnitude of the AP magnetic field scales as $B(x,y,t) = \frac{B_0(t) r_a}{\sqrt{x^2+y^2}}$ with x- and y-components $B_x(x,y,t) = B_0(t)\frac{-y}{x^2+y^2}r_a$ and $B_y(x,y,t) = B_0(t)\frac{x}{x^2+y^2}r_a$, and z-component $B_z = 0$. The average field over a square layer of area $l \times l$ and thickness d is thus

$$B_{x,\text{avg}}(t) = -B_0(t)\frac{r_a}{l^2 d}\int_{-\frac{l}{2}}^{\frac{l}{2}}dz\int_{-\frac{l}{2}}^{\frac{l}{2}}dx\int_{-r_a}^{-(r_a+d)}dy\,\frac{y}{x^2+y^2}$$
$$= B_0(t)\left(\frac{r_a}{2d}\ln\left(\frac{(r_a+d)^2+\left(\frac{l}{2}\right)^2}{{r_a}^2+\left(\frac{l}{2}\right)^2}\right)+\frac{r_a^2}{\left(\frac{l}{2}\right)d}\tan^{-1}\left(\frac{r_a}{\left(\frac{l}{2}\right)}\right)-\frac{r_a(r_a+d)}{\left(\frac{l}{2}\right)d}\tan^{-1}\left(\frac{r_a+d}{\left(\frac{l}{2}\right)}\right)+\frac{\pi r_a}{l}\right).$$

19

A similar equation exists for $B_{y,\text{avg}}(t)$. Here we have made the valid assumption that the layer dimension along z is small compared to the characteristic AP length scale, $l \ll \lambda_{\text{AP}}$, such that the magnetic signal is not attenuated when averaged over this dimension. Again, for $r_a = 5$ μm and $d = 5$ μm, and with $l = 5$ μm, the expected magnetic field averaged over the sensing region is reduced only slightly further to 0.66 nT.

Expected magnetic signals from non-axonal structures

The calculations above rely on a wire-type model to approximate neuronal activity of the giant axon of *M. infundibulum*. As discussed previously, the axial symmetry of the giant axon justifies use of this model. For activity in other neuronal processes that do not possess axial symmetry, such as the soma[40], different models can be used to predict the magnetic signal waveforms.

The geometry of the soma may be approximated to first order as an isotropic spherical cell. The spread of current in this cell model has been studied extensively[38,53,54,55] when calculating steady state membrane potentials in response to current injections. The relevant conclusions from these calculations are: (i) the time constant of responses to a change in current in the soma is up to three orders of magnitude faster than axonal membranes[40]; (ii) in steady state, the voltage distribution constitutes an isopotential with small local variations occurring near the sites of current injection; and (iii) for cell bodies that are ellipsoidal rather than spherical, the voltage distribution deviates more globally from an isopotential, as further discussed herein.

To apply these conclusions to an estimate of the magnetic field signal from a neuron cell body at an NV sensor layer, we relate the magnetic field to the dominant intracellular current through the Biot-Savart law, and we apply the same assumption used in the axon model, namely that the dominant currents are ohmic. We determine the spatial distribution of intracellular currents, and thus detected magnetic fields at the nearby NV sensor layer, using the gradients of the intracellular voltage[40].

For current injected into the center of an isotropic conducting spherical soma, the spatial voltage gradient, and thus the current flow, is directed radially outward due to spherical symmetry. Radial current sources in a spherical conducting medium do not produce a net magnetic field[56].

Next we more formally treat the case wherein the current injection point is on the surface of the soma, at $r = a$, $\theta = 0$, $\phi = 0$ in spherical coordinates, where a is the radius of the spherical soma. The spread of current from this injection point is azimuthally symmetric (independent of ϕ). For a current step from $I = 0$ to some finite value I_0 initiated at $t = 0$, the voltage response inside the soma is given by[38]

$$V_i(r,\theta,t) = \frac{I_0 R}{4\pi a^2}\sum_{n=0}^{\infty}(2n + 1)\left(\frac{r}{a}\right)^n \frac{(n+1)\sigma_e P_n(\cos\theta)\left(1-\exp\left(-\frac{t}{\tau}(1+\alpha_n)\right)\right)}{(n\sigma_i + (n+1)\sigma_e)(1+\alpha_n)},$$

20

where

$$\alpha_n \equiv \frac{Rn(n+1)}{a\left(\frac{n}{\sigma_e}+\frac{n+1}{\sigma_i}\right)}.$$

21

Here $P_n(\cos\theta)$ are the Legendre polynomials and τ is the time constant of the membrane, which can be expressed using an RC-circuit analogy as an effective resistance R times an effective capacitance C. The conductivities σ_i (internal) and σ_e (external) arise here due to the continuity boundary conditions at the surface of the membrane.

The leading term in the summation, $(n = 0)$, represents an isopotential, having no spatial dependence:

$$V_{i,n=0}(r,\theta,t) = \frac{I_0 R}{4\pi a^2}\left(1-\mathrm{e}^{-\frac{t}{\tau}}\right).$$

22

Assuming typical values for soma electrical and geometric properties[38] yields $\frac{V_{i,n=1}}{V_{i,n=0}} \sim 10^{-4}$. Thus the isopotential remains a very good approximation for nearly all spherical soma models, even if the electrical and geometrical parameters differ somewhat from the typical values for cortical neurons considered in ref. 38. The spatial dependence is present only in the much smaller, higher order terms.

We expect this picture to be valid near the center of the soma where the current profile is not significantly affected by irregular soma boundaries. On the cell body surface, however, there are expected to be detectable magnetic signals caused by deviations from spherical symmetry and uniform conductivity. For example, relatively large magnetic signals at the cell body surface are expected to originate from axon hillocks, as these structures contain a higher density of ion channels[57]. The peak magnetic field at the surface of an axon hillock should be of roughly the same order as that from an axon of the same diameter[57] (~1 nT for mammalian cells as estimated above).

Whereas the spherical conductor model presented here predicts no magnetic field at the soma surface, a wire-type model based on cable theory can provide a crude estimate of the peak magnetic field on the surface of a nonspherical soma. The soma may be treated as a cylindrical, passive structure[58] of radius r_s, fed by a source current from an active neuronal process of radius r_a with peak magnetic field B_a. Assuming current conservation at the interface between the two structures, the estimated field at the surface of the soma would be $B_s \sim B_a\, r_a/r_s \approx 200$ pT for $r_a = 2$ μm, $r_b = 10$ μm[40], and $B_a \approx 1$ nT.

More precise estimates of electrical dynamics and associated magnetic signals at the soma and other neuronal processes will require more complex geometrical models that incorporate intricacies of the neuron shape[59]. Since the spherically symmetric model predicts no magnetic field, any magnetic signal from the soma measured by a next-generation NV-diamond instrument will likely indicate structural or electrical nontrivialities of the cell body, and hence could add to the present understanding of electrical dynamics in the soma.

Additionally, spatially inhomogeneous transient currents in the soma are expected over very short timescales compared to the membrane RC time constant[55]. The rapid spread of currents and equilibration into somas has been demonstrated through voltage imaging methods[41] where, in videos of the propagation of a voltage wavefront, the soma becomes an isopotential in less than the duration of a single postprocessed frame after interpolation (~10 μs). We anticipate that with enhanced temporal resolution in a next-generation diamond magnetic imager for neurons, we could observe magnetic signals due to spatially varying transient currents within the soma, although in the near term this will likely require repeated stimulations and temporal averaging.

Furthermore, as the geometry of an axon hillock is similar to that of a compressed tapered axon, we envision that the technique and analysis presented in this work could provide insight into future studies of electrical activity of axon hillocks and ultimately aid in improving understanding of action potential initiation[60].

Tradeoff between signal-to-noise ratio, spatial resolution, and temporal resolution in next-generation magnetic imager

In the idealized case of uniform optical intensity, MW intensity, and NV density, the shot-noise-limited magnetometer sensitivity scales as $\eta \propto 1/\sqrt{V}$, where V is the total sensing volume. We denote the volume-normalized sensitivity $\eta^V \equiv \eta\sqrt{V}$. As discussed above, the present NV-diamond magnetometer has $\eta^V = 34$ nT μm$^{3/2}$Hz$^{-1/2}$; whereas a next-generation magnetic sensor is expected to achieve $\eta^V = 118$ pT μm$^{3/2}$Hz$^{-1/2}$, nearly 300 times better than the present system.

The sensing volume V consists of the NV layer at the diamond chip surface, with area A and depth d such that $V = Ad$. The sensing surface may be divided into a square grid of M pixels, with each pixel forming the top surface of a voxel of volume $v = l^2 d$, where $l = \sqrt{A/M}$ is the pixel side-length. The sensitivity for a pixel is then $\eta_{\text{pixel}} = \frac{\eta^V}{l\sqrt{d}}$. The scaling relation between η_{pixel} and l is illustrated in Fig. S13*A*.

For a shot-noise-limited magnetometer, higher temporal resolution requires increased magnetometer bandwidth BW. The SNR for a fixed amplitude magnetic signal varies as SNR $\propto 1/\sqrt{\text{BW}}$. For this exercise, the desired bandwidth is assumed to be from DC to BW. The rise time, a measure of the temporal resolution, is related to the bandwidth by BW(Hz) = $0.35/\tau_{10/90}$(s). By substitution, SNR $\propto \sqrt{\tau_{10/90}}$. Moreover, since $\eta_{\text{pixel}} = \frac{\eta^V}{l\sqrt{d}}$, the SNR for a single pixel also scales as SNR $\propto l$. Thus the SNR for a pixel remains constant when the quantity $\tau_{10/90} l^2$ is held constant. For example, to gain a factor of 4 × in temporal resolution while maintaining fixed SNR, the spatial resolution of the magnetic field map would have to be made 2 × worse. This illustrates the inherent tradeoff between spatial resolution and temporal resolution if a given SNR is desired.

For signal fields containing spatial information, however, the SNR may level off or even decrease as l increases, as discussed in more detail below, which can also affect the tradeoff between $\tau_{10/90}$ and l for fixed SNR. See Fig. S13*B* for an example estimated SNR scaling for a typical mammalian axon ($r_a = 5$ μm) with spatial resolution l, for fixed NV layer depth ($d = 5$ μm). In such cases, smaller sensing volumes are often preferable, even for fixed η^V. The quantity $1/(\eta V)$ is a useful figure of merit for imaging non-uniform fields from mammalian neurons, as this quantity encompasses the tradeoff between SNR and spatial resolution for magnetic fields that fall off as $1/\sqrt{V}$ (i.e., as $1/l$ for fixed d). In a next-generation NV-diamond magnetometer with $V \approx 50$ μm × 50 μm × 5 μm = 1.25×10^4 μm^3, the value of $1/(\eta V)$ is expected to be enhanced by ~6000 compared to the present instrument.

Optimization may be performed to find the ideal sensing volume per pixel for a given bulk sensitivity by considering the spatial information of the fields being sensed. As described above, the spatially averaged AP magnetic field from an axon of radius r_a, placed on a sensing layer of area $a = l \times l$ and thickness d, and oriented along the z-axis, is given by Eqn. 19. In the limit where $d, l \ll r_a$, the signal approaches $B_0(t)$. For $l \gg r_a, d$, the signal falls as $1/l$, and for $d \gg r_a, l$ the signal falls off as $1/d$. As shown in Fig. S13*B,* there is a tradeoff between spatial resolution and SNR for a fixed sized axon. Determination of the optimal spatial resolution requires evaluating when the marginal resolution cost of increasing l is no longer justified by increased SNR, as shown in Fig. S13*B*. For a signal field that approximates the field from a current-carrying wire as considered above (from the axon model[36,37]), and for pixel sensing volumes much larger than the wire radius but still smaller than the wire length, such that the above equation for $B_{x,\text{avg}}(t)$ continues to apply, the average magnetic field over a sensing volume $v = l^2 d$ falls off as roughly $1/l$ for fixed d, and the shot-noise limited SNR is independent of l. In this flat asymptotic limit shown in Fig. S13*B*, enhancing the spatial resolution has negligible cost to the SNR. In contrast, when l becomes small compared to r_a, the signal field becomes roughly uniform over each sensing volume. At this point, the SNR per pixel falls off with decreasing l, as shown in Fig. S13*B*, and the additional spatial information gained from the enhanced resolution diminishes. However, in the absence of camera read noise, more pixels - each corresponding to smaller sensing regions - may be used and later binned together in post-processing to recover larger SNR values.

Based on this analysis, where for simplicity we have ignored camera read noise, we conclude that the optimal regime in which to operate an NV-diamond neuron magnetic imager with fixed η^V (in the absence of image post-processing) is when both d and l are $\approx r_a$. For the example of a neuron with $r_a = 5$ μm, (and choosing $d = r_a$), Fig. S13*B* marks with an open circle the approximate crossover between the two limiting regimes at $l = 10$ μm. Use of non-square sensing regions ($a = l \times w, l \neq w$) matched roughly to the dimensions of the sources being imaged can lead to significant improvements in expected SNR, as discussed below.

Estimated signal-to-noise ratio and detection probability of a next-generation instrument

Here we estimate the SNR and event detection probability of a mammalian neuron AP magnetic signal for a next-generation NV-diamond magnetometer, which, as discussed above, has an anticipated volume-independent sensitivity of 118 pT μm$^{3/2}$Hz$^{-1/2}$. For a sensing volume of 5 μm × 12.5 μm × 200 μm, and a 5-μm-radius mammalian axon centered on and oriented along the long axis of the sensing region with a 1 nT peak AP magnetic field at the surface of the axon, the spatially-averaged magnetic field detected by the NV-diamond sensor is 564 pT. The RMS noise of a time trace of magnetic data taken over that sensing volume with sampling rate 3 kS/s is 57.9 pT. The expected SNR from this estimate is thus 9.73, prior to any SNR-enhancing temporal filtering techniques.

We next employ established methods to estimate the spike detection probability and error rate for neuron AP sensing. We equate the SNR to the discriminability parameter d' in ref. 61 and assume Gaussian distributed white noise and no prior information about expected firing rates. For $d' = 9.73$ the expected detection probability is $1 - 5.72 \times 10^{-7}$ and the expected false positive probability per sample is 5.72×10^{-7}, yielding an overall detection error probability of $2 \times 5.72 \times 10^{-7} = 1.14 \times 10^{-6}$. With 3 kS/s, the detection error rate is 0.00343 per second. However, because the NV-diamond magnetic technique can temporally resolve the AP waveform shape, the constraints on what signals are interpreted as AP events can be tightened beyond a spike threshold, significantly reducing the false detection rate, as is employed for microelectrode arrays with similar temporal resolution[62]. For example, the detected signal can be correlated to an expected waveform, similar to the matched filtering applied to data presented in this work, and signals that do not fit the expected waveform shape can be rejected. In addition, with magnetic imaging implemented, spatial correlations may be used in the same way to limit the number of false detections. Furthermore, knowledge of the spatiotemporal dynamics of APs can also be used to decrease the error detection rate, for example, by requiring signals detected as APs to propagate through the image in time. Many of these spatial and temporal filtering techniques are already established and heavily leveraged in all-optical electrophysiology methods[41].

Extension of technique to resolving currents from neuronal networks and non-axonal processes

Realization of sufficient magnetometer sensitivity is anticipated to be the primary challenge to extending NV-diamond magnetic imaging to mammalian neurons and from single neurons to networks of neurons. In contrast, implementation of magnetic imaging with micron-scale spatial resolution and ~100 μm field of view is expected to be relatively straightforward, as it mainly requires applying previously developed wide-field NV-diamond magnetic imaging technology demonstrated with biological systems[5] to the present and anticipated enhanced-sensitivity magnetometer. Sensitivity rather than spatial resolution is thus the chief limit to reconstructing neuronal currents from acquired magnetic field maps using an NV-diamond magnetic imager. In the absence of noise, the biomagnetic inverse problem[44] can be uniquely solved for 2D networks close[63] to the diamond sensing layer. 3D networks are more problematic: not only are magnetic signals smaller due to greater source-sensor standoff distance, but the problem is under-constrained, meaning that multiple reconstructions of current sources are possible from an observed magnetic field map. In all cases, *a priori* information about the network structure under study, as could be obtained with a bright field confocal microscope or other traditional optical methods, may be used to further constrain solutions, thereby increasing reconstruction accuracy and decreasing magnetometer sensitivity requirements. Similarly, prior information about axon morphology, the presence or absence of myelin, and other physiological and geometrical features could inform determination of the origins of imaged magnetic signal differences and associated individual neuronal currents, e.g., due to differences in AP conduction velocity[64]. To resolve signals from networks of neurons and non-axonal processes, we envision a staged approach outlined in more detail below. For this discussion we define the mean neuron-to-neuron spacing as w and the mean neuron-to-sensor-plane standoff distance as q. Note also that each stage of this effort will be aided by studies of model physical systems (e.g., fabricated wire arrays placed on the NV-diamond sensor surface) to characterize system performance and current source reconstruction algorithms for well-controlled systems.

The first stage consists of imaging sparse ($w > q$, where $q \sim 10$ μm), 2D arrays of neurons and reconstructing the axonal currents. To evaluate the feasibility of imaging 2D arrays of axons, a Monte Carlo simulation is performed of the magnetic signal from four axons in close proximity to each other. The simulation assumes a shot-noise-limited sensitivity of 118 pT $\mu m^{3/2} Hz^{-1/2}$, as outlined above. The field-of-view of the simulation is 100 μm × 100 μm and the NV layer depth is d = 5 μm. The pixel size is set to 2.5 μm × 2.5 μm. Axons are modeled as straight wire-like sources according to the axon (volume conductor) model[35,36] with $r_a = 5$ μm, placed in contact with the diamond NV layer; and fired simultaneously, each with action current set to produce a peak field of 1.0 nT at the axon surface. For the sensing volume corresponding to each pixel, the expected magnitude of the magnetic field vector is calculated (Fig S14*A*). A noise profile of 1.2 nT RMS deviation per pixel, calculated from expected performance of the next-generation device, is added on top of this profile to achieve the realistic profile of an image taken in a single shot with a sampling rate 3 kS/s (Fig. S14*C*).

Without substantial loss of spatial information, the SNR of the image can be improved through use of low-pass spatial filtering, as shown in Fig. S14*E*, where a simple Gaussian convolution is applied to the image in Fig. S14*C*. This spatial filter takes advantage only of the knowledge that shot-noise fluctuations have a higher spatial frequency bandwidth than the magnetic signal, as shown in Fig. S14*B, D, F*. By filtering out high spatial frequency components of the image, a Gaussian spatial filter eliminates more noise than signal from the image. This and other straightforward spatial filters are able to enhance image SNR without knowledge of the propagation direction or conduction velocity of APs. When the temporal profile of expected magnetic signals is known, temporal matched filtering can improve image SNR even in the absence of information on the spatial distribution or propagation directions of current sources. As shown in Fig. S4, the temporal profile of an axonal AP magnetic signal waveform is independent of axon orientation, even when only a single component of the magnetic field is measured. Further improved performance can be achieved using prior information about the spatial locations of the neuronal current sources (i.e. spatial matched filtering), through combining spatial information with the known temporal profile of the action current, and through use of similar image-processing methods[41]. Where possible, synchronizing multiple AP firings and averaging traces together, as is done in the present studies, can significantly improve image SNR. To represent the effect of modest ($N_{\mathrm{avg}} = 10$ trials) averaging, the noise level is decreased to 1.2 nT/$\sqrt{10}$ = 0.35 nT in Fig. S14*G* to demonstrate the distinguishability of the four axons in the image, even from the raw unprocessed magnetometer output.

Furthermore, methods of spatial filtering can be used to solve the biomagnetic inverse problem — i.e., relating magnetic field maps back to current distributions — even in the absence of prior knowledge about the nature of the sources. Substantial work in the literature, (particularly by Wikswo and collaborators) demonstrates the ability to reconstruct arbitrary 1D and 2D current distributions from magnetic field images recorded some distance from the sources in the presence of noise[63,65,42]. Additionally, spatial filtering introduces a tradeoff between required SNR to determine the current distribution and standoff distance q from the imaging plane to the source plane, favoring small q. For example, ref. 63 shows that two magnetic imagers — one with 100 × the noise and 1/5 the SNR of the other, but positioned 3× closer to the source — can reconstruct a 2D current distribution equally well. Growth of 2D cultures of neurons directly on the diamond sensor alongside use of a sufficiently thin NV sensor layer ($d \sim 5$ μm) will ensure that q is always small compared to both w and λ_{AP}. Thus, we envision assigning magnetic fields to particular cells in sparse networks, without prior knowledge of the nature of the signals, based on magnetic field maps with limited SNR.

The next stage involves magnetic imaging of currents from dense 2D neuronal networks comprising overlapping neurons with mean neuron-neuron spacing $w \lesssim 10$ μm. To avoid aliasing of high spatial frequencies when using spatial filtering methods, $w \gtrsim l, d$ is generally required[63,66]. Fortunately, the sensing surface of an ensemble-based NV imager is continuous, with imaging resolution adjustable external to the sensor (e.g., by changing imaging magnification), with the restriction that the ultimate resolution will be limited to approximately the sensor layer thickness d. With sufficient magnetometer sensitivity, spatial filtering techniques are also expected to allow effective reconstruction of currents in non-axonal processes in 2D, even without *a priori* knowledge of the currents within such structures. Furthermore, super-resolution imaging techniques could be applied to the diamond sensor to enhance spatial resolution below the optical diffraction limit[67]. Therefore, the principle challenge in magnetic imaging of dense 2D neuronal networks reduces to improving magnetometer sensitivity.

Assuming successful implementation of the above, we next foresee reconstruction of sparse neural networks in 3D space, e.g., cultures or tissue slices with nonnegligible thickness (~100 μm) placed on the diamond surface. We expect to maintain a proximity of ~10 μm from the sensor layer to the tissue slice, although the standoff distance q to a given current source within the slice will

naturally be larger (up to ~100 μm), reducing magnetic signal amplitudes and limiting the effective spatial resolution to ~q in the absence of additional known structural information localizing specific current sources within the tissue.

Solving the inverse problem in 3D requires *a priori* assumptions about the nature of the expected sources in order to determine a unique solution[63]. For example, schemes used in present-generation MEG typically employ a dipole model of currents in the brain generating detectable magnetic fields[56]. While this simplification sufficiently constrains the inverse problem and allows for physical solutions to be found, it largely neglects the complex physiological activity generating the fields — effects that often cannot be observed at distances far from the sources such as outside the skull. Furthermore, in some cases dipole models are physiologically unrealistic[42].

Where possible, it is preferable to use a model that retains information on individual neurons. For example, the axon (volume conductor) model[35,36], which constitutes the basis for the simple magnetic model used in the present work, treats individual neurons as cylindrically symmetric wire-like current sources. In order to resolve currents from individual neurons and justify use of this model, magnetic field measurements must be made at distances sufficiently close to the neurons, such that $q \lesssim w$[42]. For SQUID and atomic magnetometers, both of which are bulky and typically cannot meet this requirement, the method of downward continuation[43] may be used to reconstruct the field distribution in the near-field[42]. However, downward continuation requires a much improved SNR and places additional restrictions on both acceptable values of w and the nature of the sources[43]. The close proximity to sources afforded by the NV-diamond sensor greatly alleviates these restrictions[63]. In the near-field regime, ($\boldsymbol{q < \lambda_{AP} \approx 1}$ mm), magnetic fields from simultaneous currents at different locations will neither partially nor completely cancel as they would at a location ~1 cm or more away[42]. The ability of NV-diamond to perform magnetic field measurements in the near-field, where return-currents can largely be ignored, allows for the application of more physically detailed, physiologically realistic models to sufficiently constrain the inverse problem without ignoring physiologically relevant information about the neuronal dynamics. Furthermore, the high temporal resolution of NV magnetometry can allow for nearly-overlapping current sources to be resolved as long as their signals are separated in time. To further aid in simplifying the substantial inverse problem posed by dense 2D and sparse 3D networks, cultures may be encouraged to grow in controlled patterns[68,69].

Resolving activity from non-axonal processes in *dense* 3D networks with both high spatial and high temporal resolution is a significant challenge in neuroscience[70], which we regard as a longer-term goal that may realistically require additional structural knowledge of the network, such as provided by a confocal scan, along with further sensitivity enhancements of the magnetic imaging technique. We regard the task of achieving the necessary sensitivity to map neuronal network dynamics as a practical challenge, albeit a difficult one, which is limited more by a need for engineering advances in NV-diamond magnetometry than by fundamental constraints.

1. Binstock L, Goldman L (1969) Current and voltage-clamped studies on Myxicola giant axons: Effect of tetrodotoxin. *J Gen Physiol* **54**:730–740.
2. Rudy B (1981) Inactivation in Myxicola giant axons responsible for slow and accumulative adaptation phenomena. *J Physiol* **312**:531–549.
3. Hines M L, Carnevale N T (1997). The NEURON simulation environment. *Neural Comput* **9**:1179-1209.
4. Nicol J A C (1948) The giant nerve-fibres in the central nervous system of Myxicola (Polychaeta Sabellidae). *Q J Microsc Sci* **89**:1–45.
5. Carpenter D, Hovey M, Bak A (1975) Resistivity of axoplasm. II. Internal resistivity of giant axons of squid and Myxicola. *J Gen Physiol* **66**:139–148.
6. Doherty M W, *et al.* (2013) The nitrogen-vacancy colour centre in diamond. *Phys Rep* **528**:1–45.
7. Budker D, Romalis M (2007) Optical magnetometry. *Nat Phys* **3**:227–234.
8. Pham L M, *et al.* (2012) Enhanced metrology using preferential orientation of nitrogen-vacancy centers in diamond. *Phys Rev B* **86**:121202.
9. Taylor J M, *et al.* (2008) High-sensitivity diamond magnetometer with nanoscale resolution. *Nat Phys* **4**:810–816.
10. Pham L M, *et al.* (2011) Magnetic field imaging with nitrogen-vacancy ensembles. *New J Phys* **13**:045021.
11. Le Sage D, *et al.* (2013) Optical magnetic imaging of living cells. *Nature* **496**:486–489.
12. Vanier J, Audoin C (1989) The Quantum Physics of Atomic Frequency Standards. (A Hilger).
13. Wolf T, *et al.* (2015) Subpicotesla diamond magnetometry. *Phys Rev X* **5**:041001.
14. Nicol J A C (1948) The giant axons of annelids. *Q Rev Biol* **23**:291–323.
15. Schoenfeld R S, Harneit W (2011) Real time magnetic field sensing and imaging using a single spin in diamond. *Phys Rev Lett* **106**:030802.
16. Dréau A, *et al.* (2011) Avoiding power broadening in optically detected magnetic resonance of single NV defects for enhanced dc magnetic field sensitivity. *Phys Rev B* **84**:195–204.
17. Jensen K, *et al.* (2014) Cavity-enhanced room-temperature magnetometry using absorption by nitrogen-vacancy centers in diamond. *Phys Rev Lett* **112**:160802.
18. Shin C S, *et al.* (2012) Room-temperature operation of a radiofrequency diamond magnetometer near the shot-noise limit. *J Appl Phys* **112**.
19. Popa I, *et al.* (2004) Energy levels and decoherence properties of single electron and nuclear spins in a defect center in diamond. *Phys Rev B* **70**:201203.
20. Acosta V M, Jarmola A, Bauch E, Budker D (2010) Optical properties of the nitrogen-vacancy singlet levels in diamond. *Phys Rev B* **82**:201202.
21. Turin G (1960) An introduction to matched filters. *IRE Trans Inform Theory* **6**:311-329.
22. Roberts M B V (1962) The rapid response of Myxicola infundibulum. *J Mar Biol Ass UK* **42**:527-539
23. Shapiro M G, Priest M F, Siefel P H, Bezanilla F (2013) Thermal mechanisms of millimeter wave stimulation of excitable cells. *Biophysics J* **104**: 2622-2628.
24. Wikswo J P, Barach J P (1980) An estimate of the steady magnetic field strength required to influence nerve conduction. *IEEE Trans Biomed Eng* **27**:722-723.
25. FDA (2014) Criteria for Significant Risk Investigations of Magnetic Resonance Diagnostic Devices.
26. Pham L M (2013) Magnetic Field Sensing with Nitrogen-Vacancy Color Centers in Diamond. PhD thesis (Harvard University, Cambridge, MA).
27. Shields B J, Unterreithmeier Q P, de Leon N P, Park H, Lukin M D (2015) Efficient readout of a single spin state in diamond via spin-to-charge conversion. *Phys Rev Lett* **114**:136402.
28. Dobrovitski V V, Feiguin A E, Awschalom D D, Hanson R (2008) Decoherence dynamics of a single spin versus spin ensemble. *Phys Rev B* **77**.
29. Le Sage, *et al.* (2012) Efficient photon detection from color centers in a diamond optical waveguide. *Phys Rev B* **85**
30. Fang K, *et al.* (2013) High-sensitivity magnetometry based on quantum beats in diamond nitrogen-vacancy centers. *Phys Rev Lett* **110**:130802.
31. Mamin H J, *et al.* (2014) Multipulse double-quantum magnetometry with near-surface nitrogen-vacancy centers. *Phys Rev Lett* **113**:030803.
32. Jiang L, *et al.* (2009) Repetitive readout of a single electronic spin via quantum logic with nuclear spin ancillae. *Science* **326:**267-272.
33. Neumann P, *et al*. (2010) Quantum register based on coupled electron spins in a room-temperature solid. *Nat Phys* **6**:249-253.
34. Zhang H, *et al.* (2014) Silicon-vacancy color centers in nanodiamonds: cathodoluminescence imaging markers in the near infrared. *Small* **10**:1908–1913.
35. Woosley J K, Roth B J, Wikswo J P (1985) The magnetic field of a single axon: A volume conductor model. *Math Biosci* **76**:1–36.
36. Swinney K R, Wikswo J P (1980) A calculation of the magnetic field of a nerve action potential. *Biophys J* **32**:719-732.
37. Roth B J, Wikswo J P (1985) The magnetic field of a single axon. A comparison of theory and experiment. *Biophys J* **48**:93–109.
38. Hellerstein D (1968) Passive membrane potentials. A generalization of the theory of electronus. *Biophys J* **8**:358-379.
39. Zhang X, Zhu S, He B (2010) Imaging Electric Properties of Biological Tissues by RF Field Mapping in MRI. *IEEE Trans Med Imaging* **29:**474-481.
40. Jack J J B, Noble D, Tsien R W (1975) Excitable current flow in excitable cells (Clarendon Press).
41. Hochbaum D R, *et al.* (2014) All-optical electrophysiology in mammalian neurons using engineered microbial rhodopsins. *Nat Methods* **11**:825–833.
42. Tan S, Roth B J, Wikswo J P (1990) The magnetic field of cortical sources the application of a spatial filtering modeling to the forward and inverse problems. *Electroencephalogr Clin Neurophysiol* **76**:73-85.
43. Parker R L (1977) Understanding inverse theory. *Ann Rev Earth Planet Sci* **5**:35.
44. Sarvas J (1987) Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Phys Med Biol* **32**:11.
45. Shelton D P (1985) Membrane resistivity estimated for the purkinje neuron by means of a passive computer model. *Neuroscience* **14**:111-131.
46. Roth A, Hausser M (2001) Compartment models of rat cerebellar Purkinje cells based on simultaneous somatic and dendritic patch-clamp recordings. *J Physiol* **535**:445-472.
47. Vetter P, Roth A, Hausser M (2000) Propagation of action potentials in dendrites depends on dendritic morphology. *J Neurophysiol* **85**:929-937.
48. Stuart G, Hausser M (1994) Initiation and spread of sodium action potentials in cerebellar purkinje cells. *Neuron* **13**:703-

712.
49. Hourez R, *et al.* (2011) Aminopyridines correct early dysfunction and delay neurodegeneration in a mouse model of spinocerebellar ataxia type 1. *J Neurosci* **31**:11795-11807.
50. Khaliq Z, Raman M (2006) Relative contributions of axonal and somatic Na channels to action potential initiation in cerebellar purkinje neurons. *J Neurosci* **26**:1935-1944.
51. Rogers J, Zornetzer S F, Bloom F E (1981) Senescent pathology of cerebellum: Purkinje neurons and their parallel fiber afferents. *Neurobiol Aging* **2**:15-25.
52. Cotman C W & McGaugh J L (1980) *Behavioral Neuroscience: An Introduction* (Academic Press, Cambridge, MA).
53. Eisenberg R S, Johnson E A (1970) Three-dimensional electrical field problems in physiology. *Prog Biophys Mol Biol* **20:**1-65.
54. Pickard W F (1971) The spatial variation of plasmalemma potential in a spherical cell polarized by a small current source. *Math Biosci* **10**:307-328.
55. Eisenberg R S, Engel E (1970) The spatial variation of membrane potential near a small source of current in a spherical cell. *J Gen Physiol* **55**: 736-757.
56. Hämäläinen M, Hari R, Ilmoniemi R, J Knuutila J, Lounasmaa O V (1993) Magnetoencephalography–theory instrumentation and applications to noninvasive studies of the working human brain. *Rev Mod Phys* **65**:413–497.
57. Wollner D A, Catterall W A (1986) Localization of sodium channels in axon hillocks and initial segments of retinal ganglion cells. *Proc Natl Acad Sci USA* **83:**8424-8428.
58. Murakami S, Zhang T, Hirose A, Okada Y (2002) Physiological origins of evoked magnetic fields and extracellular field potentials produced by guinea-pig CA3 hippocampal slices. *J Physiol* **544.1**:237-251.
59. Sakatani S, Hirose A (2003) The influence of neuron shape changes on the firing characteristics. *Neurocomputing* **52-54:**355-362.
60. Clark B D, Goldberg E M, Rudy B (2009) Electrogenic Tuning of the Axon Initial Segment. *Neuroscientist* **15**:651-668.
61. Wilt B A, Fitzgerald J E, & Schnitzer M J (2013) Photon shot noise limits on optical detection of neuronal spikes and estimation of spike timing. *Biophys J* **104**:51–62.
62. Muthmann J, *et al.* (2015) Spike detection for large neural populations using high density multielectrode arrays. *Front Neuroinform* **18**.
63. Roth B J, Sepulveda N G, Wikswo J P (1989) Using a magnetometer to image a two-dimensional current distribution. *J Appl Physics* **65**:361-372.
64. Murakami S, Okada Y (2015) Invariance in current dipole moment density across brain structures and species: Physiological constraint for neuroimaging. *NeuroImage* **111**:49-58.
65. Wikswo J P, Roth B J (1988) Magnetic determination of the spatial extent of a single cortical current source: a theoretical analysis. *Electroencephalogr Clin Neurophysiol* **69**:266-276.
66. Ahonen A L, *et al*. (1993) Sampling theory for neuromagnetic detector arrays. *IEEE Trans Biomed Eng* **40**:859-869.
67. Arai K, *et al.* (2015) Fourier magnetic imaging with nanoscale resolution and compressed sensing speed-up using electronic spins in diamond. *Nat Nanotechnol* **10**:859–864.
68. Chandrasekar I, Huettner J E, Turney S G, Bridgman P C (2013) Myosin II regulates activity dependent compensatory endocytosis at central synapses. *J Neurosci* **33**:16131–16145.
69. Turney S G, *et al.* (2015) Nerve growth factor stimulates axon outgrowth through negative regulation of growth cone actomyosin restraint of microtubule advance. *Mol Biol Cell* **26**.
70. Marblestone A H, *et al.* (2013) Physical principles for scalable neural recording. *Front Comput Neurosci* **7**.

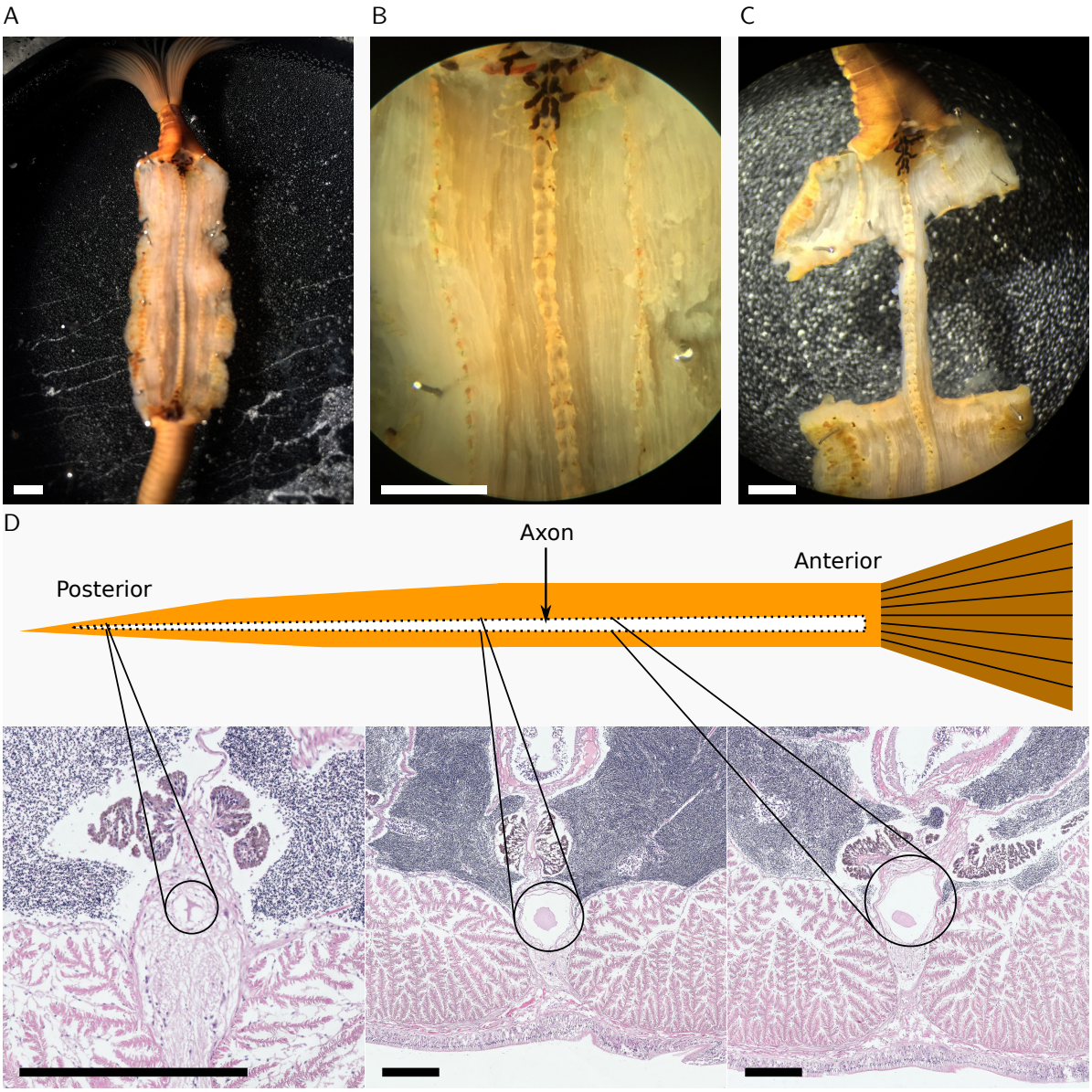

Fig. S1 | Dissected specimen of *M. infundibulum*. (*A*) Photo of worm with nerve cord containing giant axon exposed as discussed in Methods. (*B*) Close-up view of same specimen. The nerve cord is ~ 400 μm across near the anterior (top) end. (*C*) Same specimen with muscle tissue removed. At this stage the dissection protocol is complete. All white scale bars correspond to 2 mm. The levels of each photo are slightly and uniformly adjusted for improved contrast. (*D*) Cartoon drawing of worm and transverse sections. Middle and right sections are reproduced from Fig. 4*A*, *B*; leftmost section is from near the tip of the posterior end of the specimen, further demonstrating the significant tapering of the giant axon. All black scale bars correspond to 400 μm.

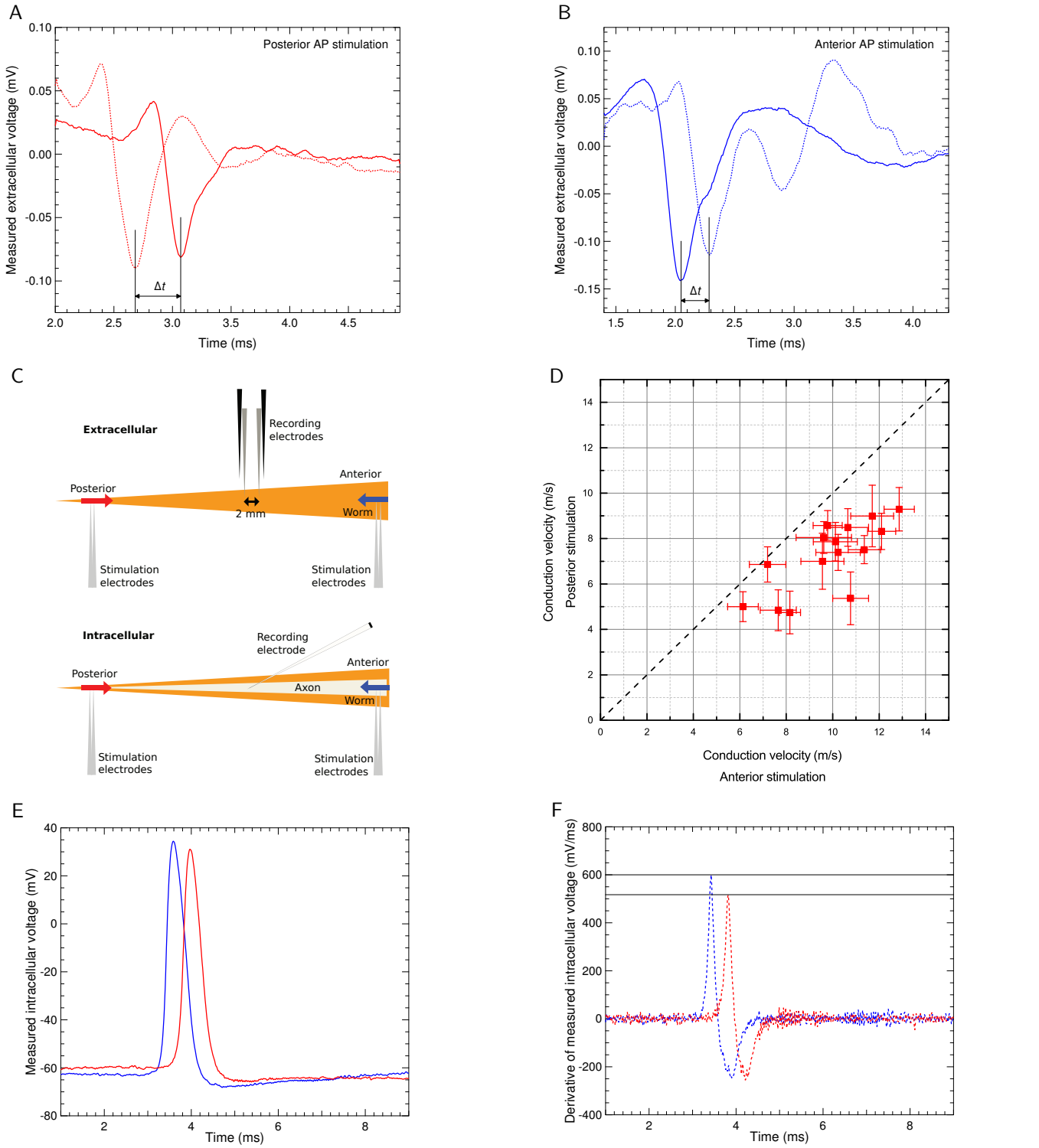

Fig. S2 | Electrophysiology experiments on asymmetry of AP in *M. infundibulum*. Representative pair of extracellular recordings of $\Phi_{ex}^{meas}(t)$ stimulated from (*A*) the posterior end (red) and (*B*) the anterior end (blue) of intact specimen of *M. infundibulum*. The time difference Δt between extrema (minima) of recorded voltage traces at two points near the mid-point of the worm spaced by 2 mm is extracted; and on average a larger Δt is found for posterior stimulation than for anterior stimulation, as shown in (*D*). For both (*A*) and (*B*) the dotted (solid) trace shows the signal from the recording electrodes positioned closer to the posterior (anterior) end of the worm. For both stimulations positions, v_c is extracted as (2 mm)/Δt. (*C*) Cartoon of experimental setup of extracellular and intracellular recordings for measuring v_c and $\Phi_{in}^{meas}(t)$ respectively. APs are stimulated from the posterior (anterior) end using bipolar electrodes placed within 2 mm of the end of the organism's tail (head) and all recordings are performed near the mid-point of the worm. Top: Two pairs of extracellular bipolar electrodes record AP signals as they propagate. Bottom: A glass micropipette electrode records the intracellular AP profile. (*D*) Average value of v_c measured for posterior stimulation ($v_{c,p}$) plotted versus v_c for anterior stimulation ($v_{c,a}$) for $n = 15$ worms with 100 AP recordings per stimulation position per worm. Error bars denote s.d. The ratio $v_{c,p}/v_{c,a} = 0.72 \pm 0.11$ (mean $\pm$ s.d.) differs significantly from unity (represented by dashed line). (*E*) Representative intracellular recordings of $\Phi_{in}^{meas}(t)$ from excised giant axon of *M. infundibulum* for posterior (red) and anterior (blue) stimulation. (*F*) Calculated time derivative of $\Phi_{in}^{meas}(t)$ for data shown in (*E*). Horizontal lines mark the maxima $\frac{\partial \Phi_p}{\partial t}$ (peak of red trace) and $\frac{\partial \Phi_a}{\partial t}$ (peak of blue trace).

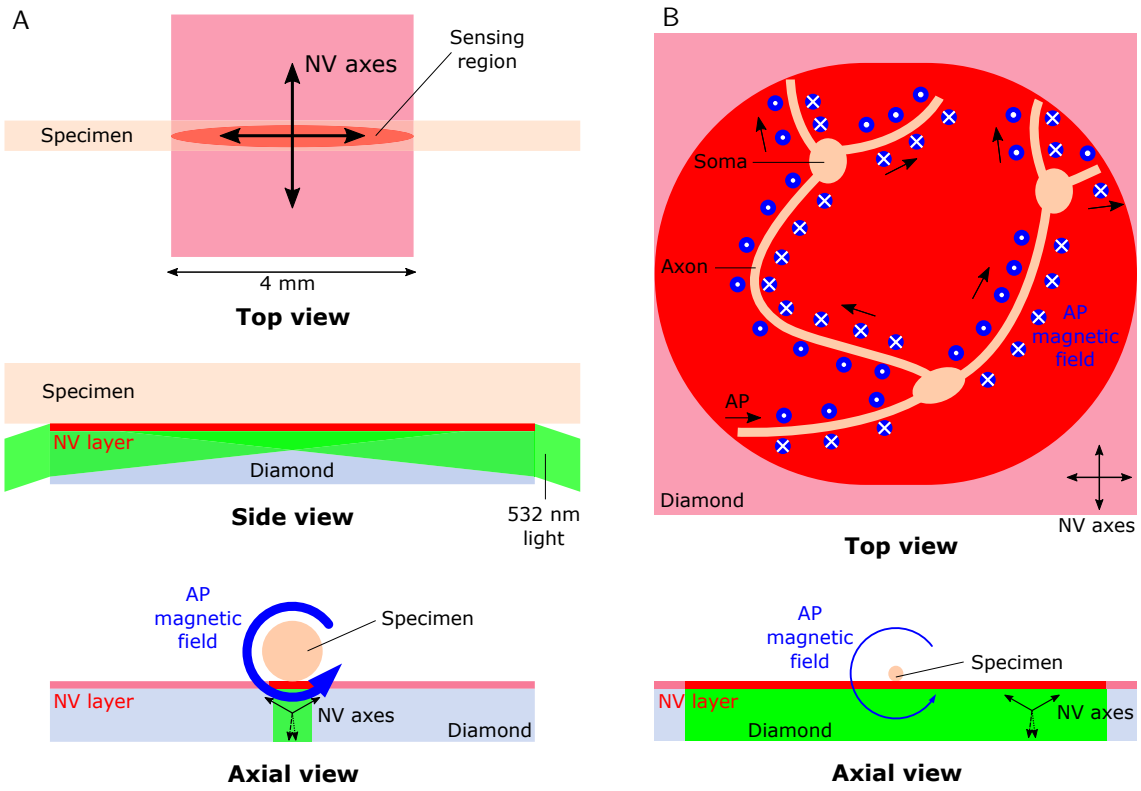

Fig. S3 | Specimen orientation with respect to NV axes. (*A*) Present specimen orientation as discussed in SI Appendix and main text. Diagram is reproduced from Fig. 1 for comparison. (*B*) Proposed method for magnetic imaging of AP dynamics from networks of smaller neurons with arbitrary orientation. Here the NV sensor layer detects the magnetic field component normal to the diamond surface, which has opposite sign on different sides of the specimen.

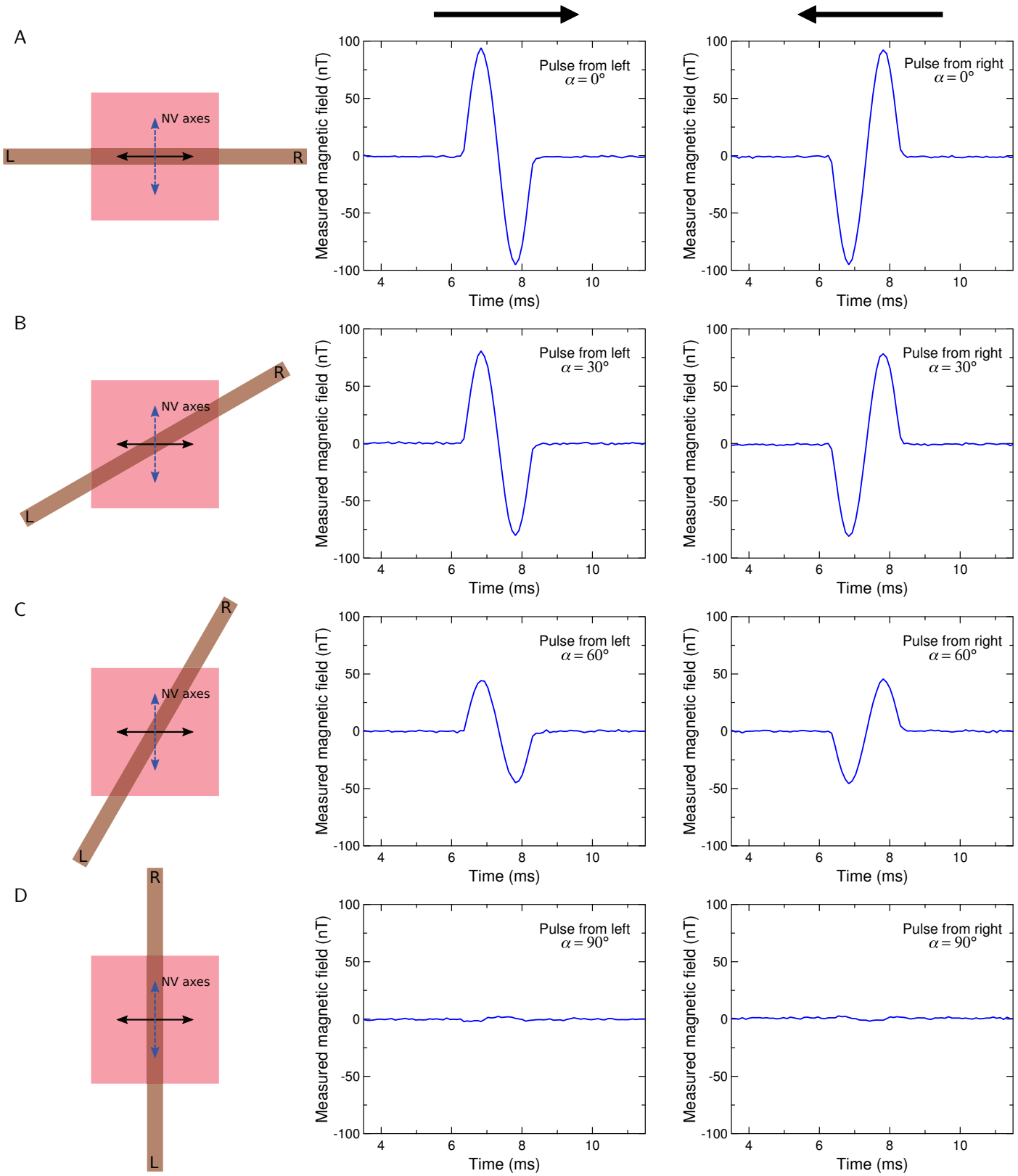

Fig. S4 | Vector magnetometry. Measured magnetic fields from biphasic current pulses (with positive polarity first) sent through a copper wire positioned on the NV-diamond surface at angles (*A*) $\alpha = 0°$, (*B*) $\alpha = 30°$, (*C*) $\alpha = 60°$, and (*D*) $\alpha = 90°$ with respect to the angle of maximum sensitivity (defined to be $\alpha = 0°$). For each angle: cartoons (left column) give the orientation of the wire on the diamond surface, where the blue dashed lines indicate the two NV axes used for sensing; and data is shown for current pulses initiated from the left (middle column) and from the right (right column). For a current pulse producing a magnetic field at the NV sensing layer of $B(t)$ = 94 nT perpendicular to the wire, the detected signal $B_x(t)$ is shown to depend on α as $B_x(t) = B(t)\cos(\alpha)$. For all angles except $\alpha = 90°$, when $B_x(t)$ goes to zero, the projected direction of pulse propagation along the x-axis (indicated by black arrows in left column cartoons) can be determined from the polarity of the magnetic signal. Each magnetic trace is an average of $N_{\mathrm{avg}} = 32$ trials.

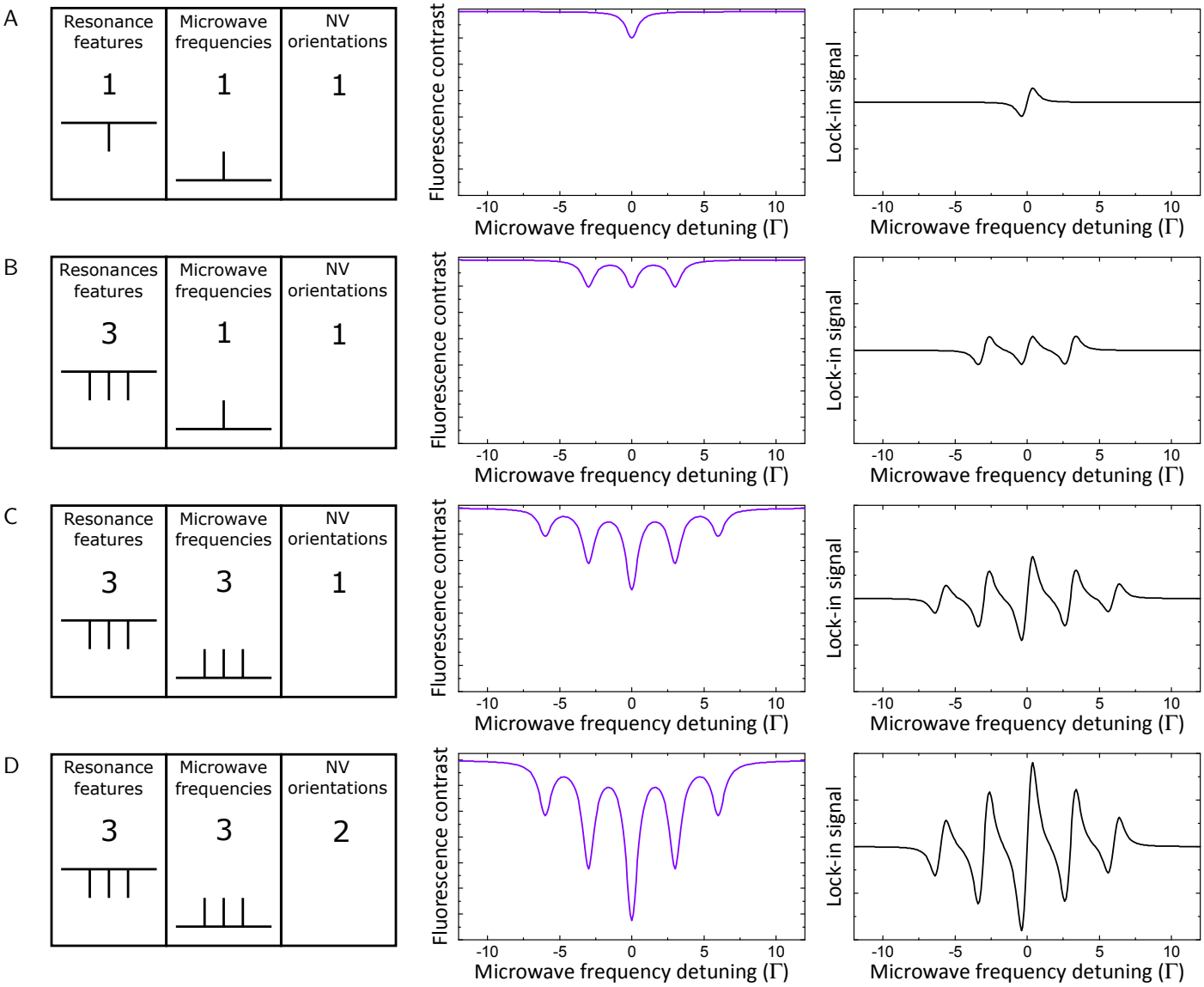

Fig. S5 | Illustration of magnetometry technique. Left column shows schematic diagrams illustrating number of ODMR features, number of MW frequencies applied, and number of NV axes used for sensing; middle column shows calculated ODMR fluorescence profiles in units of the natural linewidth Γ, in the absence of power broadening; and right column shows associated dispersion-type lock-in amplifier (LIA) signals. Fluorescence and LIA signals are given in arbitrary units. See SI Appendix for discussion of the lock-in scheme. (*A*) Diagram, fluorescence signal, and LIA signal for a single ODMR feature addressed by a single (modulated) MW frequency, sensed along a single NV axis. (*B*) Diagram, fluorescence signal, and LIA signal for three ODMR features addressed by a single (modulated) MW frequency, sensed along a single NV axis. (*C*) Diagram, fluorescence signal, and LIA signal for three ODMR features addressed by three (modulated) MW frequencies with equivalent spacing, sensed along a single NV axis. The central feature corresponds to all three applied frequencies resonantly addressing ODMR features, as described in SI Appendix. (*D*) Diagram, fluorescence signal, and LIA signal for same scenario as in (*C*) but with B_0 oriented to have equal projection along two NV axes, overlapping their ODMR features, as discussed in SI Appendix.

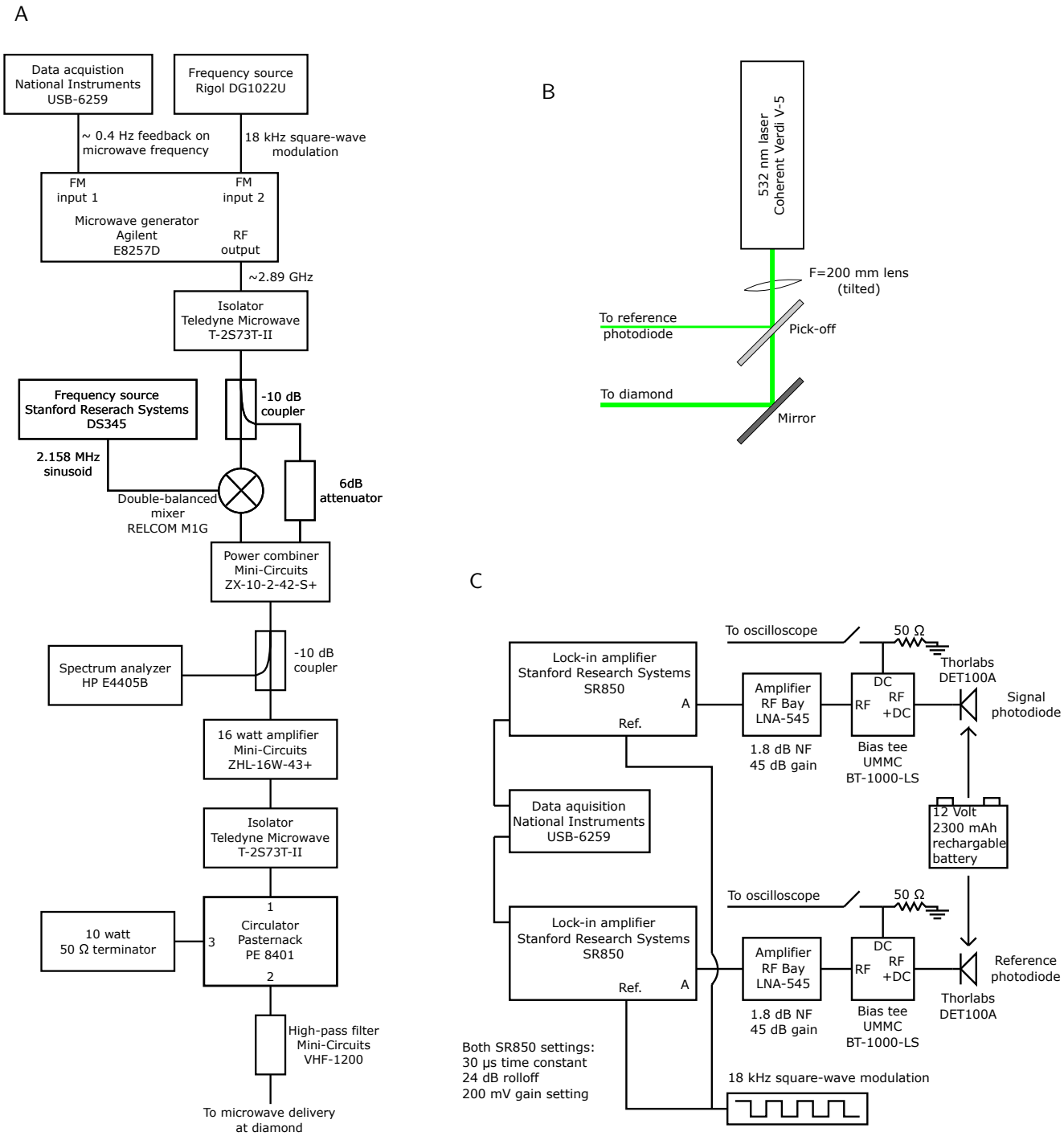

Fig. S6 | Microwave, laser, and light collection setup. (*A*) MW generation, modulation, and delivery setup as described in SI Appendix. (*B*) Laser setup as described in SI Appendix. (*C*) Signal photodiode, reference photodiode, and downstream electronics and LIAs as described in SI Appendix.

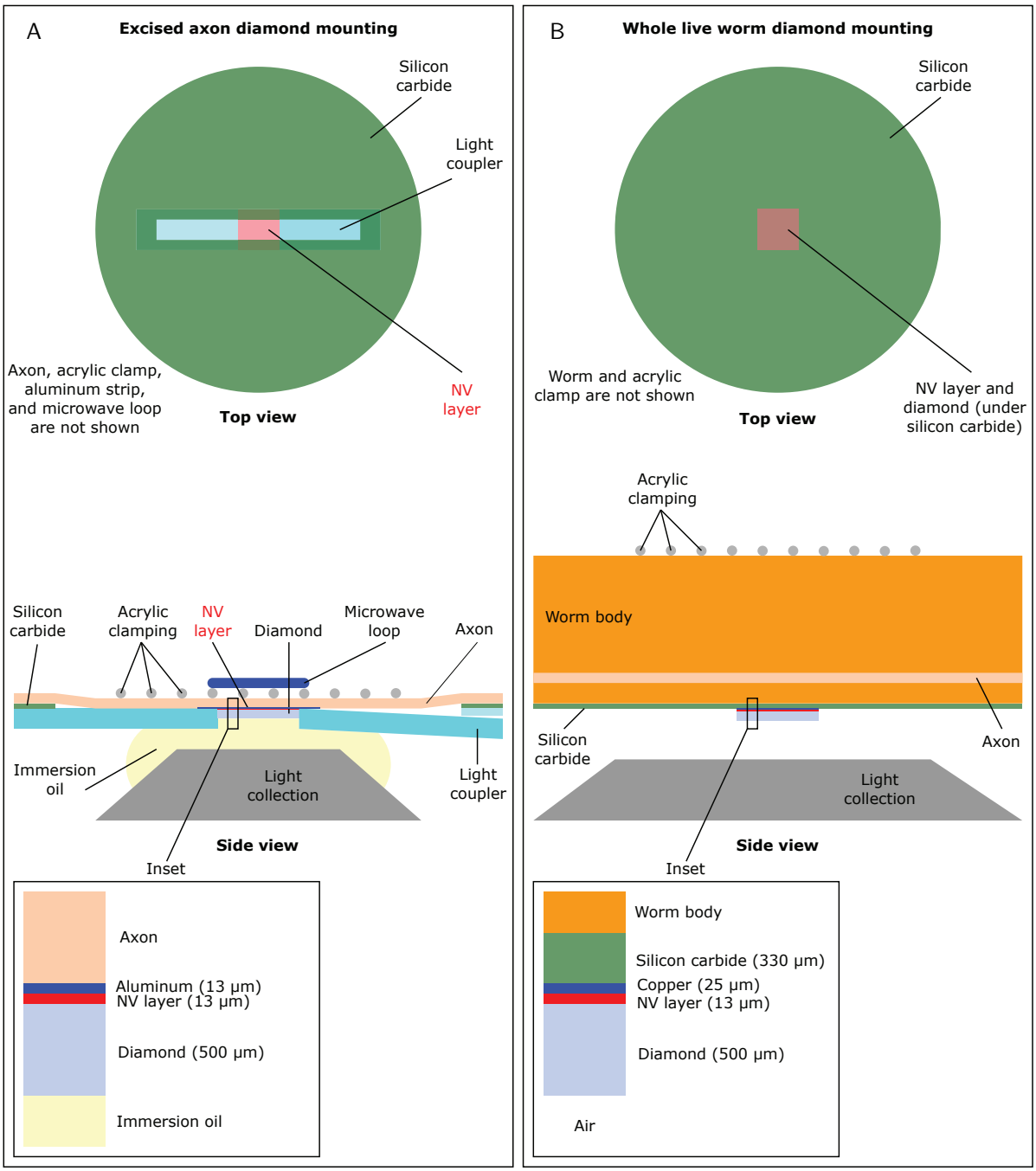

Fig. S7 | Diamond mounting. (*A*) Diamond mounting for excised axon experiments as described in SI Appendix and main text. (*B*) Diamond mounting for live, intact worm experiments as described in SI Appendix and main text. MWs are applied through the 25 μm thick copper layer.

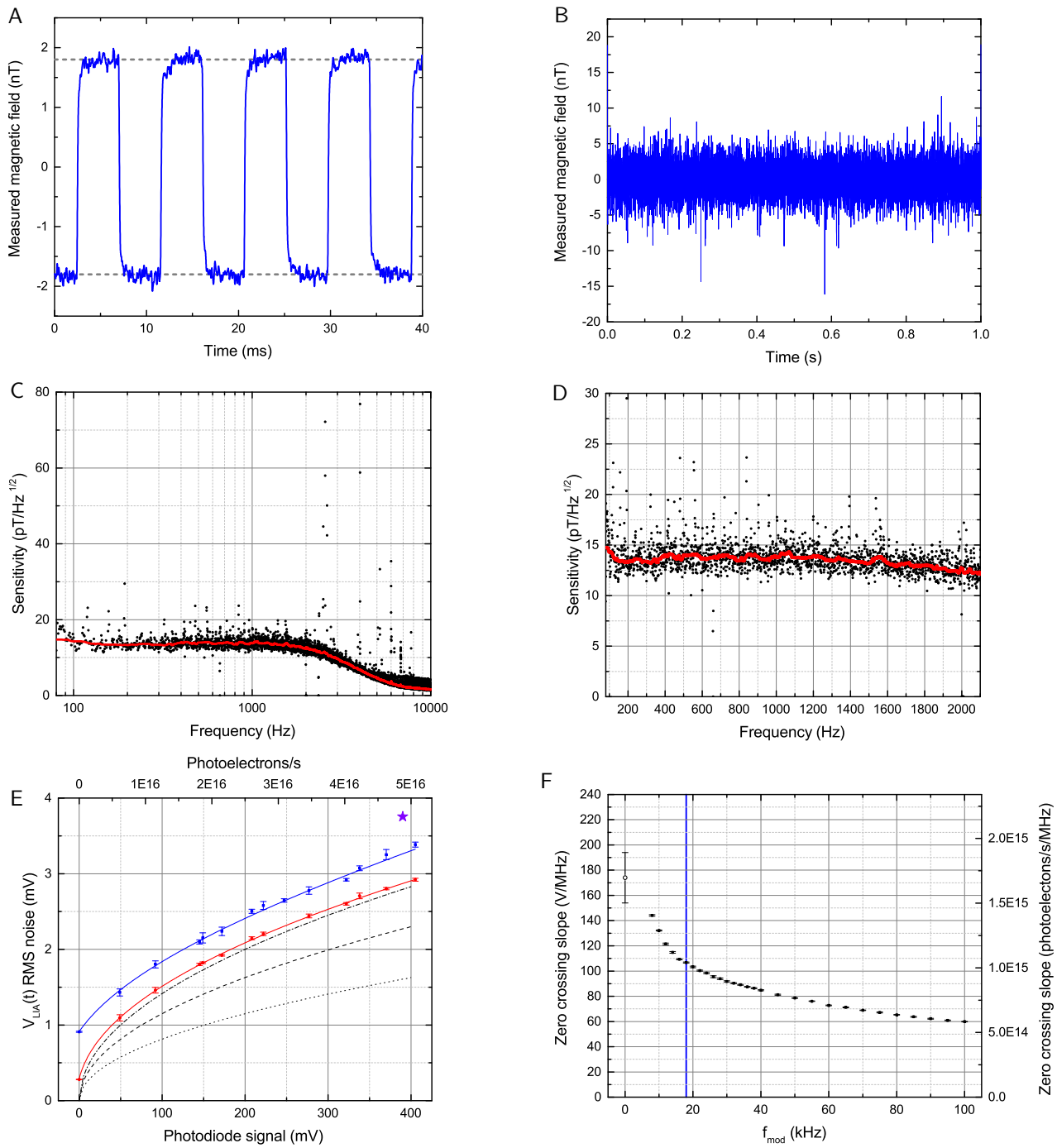

Fig. S8 | Magnetometer calibration and sensitivity. (*A*) Calibration verification as described in detail in SI Appendix. A 110 Hz square wave with 1.8 nT amplitude (calculated from coil geometry, coil distance to diamond sensor, and current through coil only) is averaged for $N_{\mathrm{avg}} = 1000$ trials. The measured magnetic field, calibrated only from the value of C_{LIA} and the lock-in amplifier voltage time trace $V_{\mathrm{LIA}}(t)$, is consistent with a 1.8 nT amplitude square wave to better than 5%. Gray dashed lines depict -1.8 nT and +1.8 nT levels. The slight rounding of the square wave's corners results from coil non-idealities rather than the magnetometer. (*B*) Measured real-time magnetic field time trace of measured magnetic field $B^{\mathrm{meas}}(t)$ with no external time-varying magnetic field applied. (*C*) Fourier transform (black points) of $N_{\mathrm{avg}} = 150$ traces of (*B*) is smoothed (red line) for clarity and is consistent with an overall sensitivity of 15 pT/$\sqrt{\mathrm{Hz}}$. (*D*) Reproduction of (*C*) with linear scale over approximate neuron signal bandwidth (80 Hz to 2 kHz). All data are taken for standard conditions ($f_{\mathrm{mod}} = 18$ kHz, nominal $\tau_{\mathrm{LIA}} = 30$ µs, with 24 dB/octave roll-off). (*E*) Measured and calculated RMS noise on $V_{\mathrm{LIA}}(t)$ versus signal photodiode voltage. Data shown by blue squares (red dots) are taken without applied MWs at $f_{\mathrm{mod}} = 18$ (90) kHz. Error bars denote s.d. Blue and red curves are fits to the respective data sets (discussed in SI Appendix), demonstrating the square-root dependence of the measured noise. Purple star marks measured noise during typical operating conditions in the presence of applied MWs. Black curves indicate calculated theoretical noise for shot noise from only the signal channel (dotted), shot noise including both the signal and reference channels (dashed), and expected noise level including shot noise from both channels and the LNA-545 amplifier noise figure of 1.8 (dot-dashed). (*F*) Measured slope of zero-crossing $\frac{dV_{\mathrm{LIA}}}{df}\Big|_{V_{\mathrm{LIA}}=0}$ with modulation frequency f_{mod}. Blue line denotes $f_{\mathrm{mod}} = 18$ kHz. Error bars denote s.d. Open circle marks the slope in the absence of modulation, calculated from the measured DC photodiode signal and the LNA-545 amplifier and LIA gains.

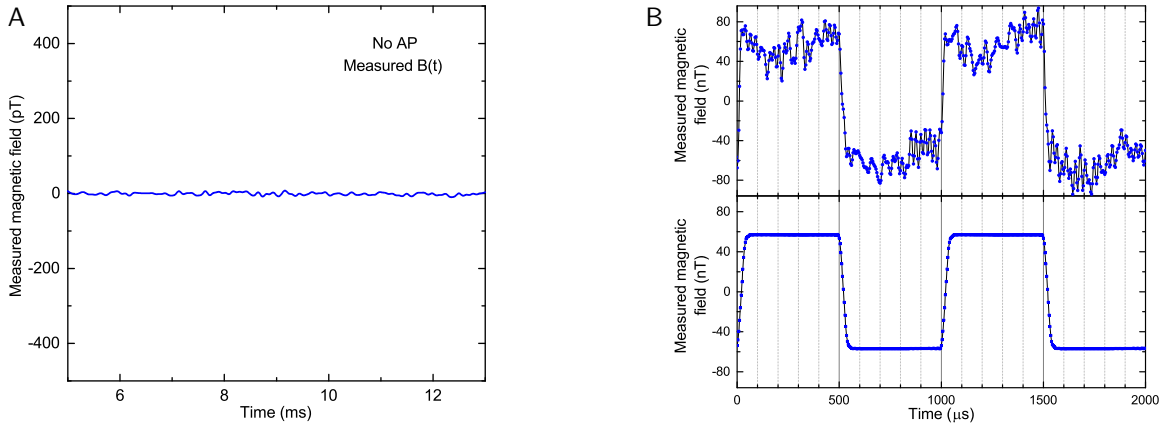

Fig. S9 | Magnetometer noise analysis and temporal resolution. (*A*) Magnetometer signal $B(t)$ under experimental running conditions in the absence of a specimen, averaged for $N_{\mathrm{avg}} = 75{,}000$ traces. No discernable pattern in the noise is present, suggesting that the magnetometer is free from correlated noise. (*B*) Top: Measured real-time magnetic field time trace of 1 kHz square wave with 57 nT amplitude as described in SI Appendix. Bottom: Same experimental setup as above, but highly averaged ($N_{\mathrm{avg}} \sim 10^6$). Data analysis indicates a 10% - 90% rise time of $\tau_{10/90} = 32$ μs. For these data only, $f_{\mathrm{mod}} = 60$ kHz, $\tau_{\mathrm{LIA}} = 10$ μs nominally, and a 6 dB/octave roll-off is used, yielding a measured $f_{\mathrm{ENBW}} = 33$ kHz. Data are FFT low-pass filtered at 45 kHz.

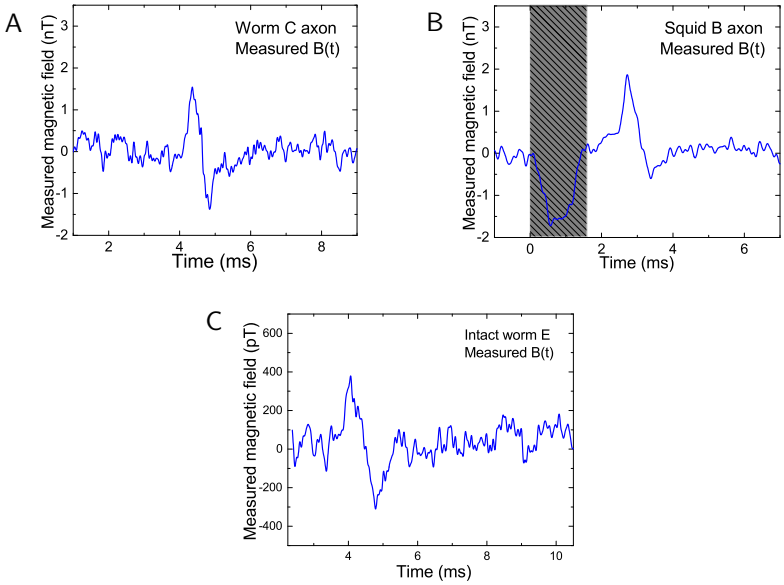

Fig. S10 | Biological replicates and extended duration sensing. (*A*) Biological replicate of data in Fig. 2*C*, except with $N_{\mathrm{avg}} = 150$. (*B*) Biological replicate of data in Fig. 2*D*. (*C*) Measured time trace of AP magnetic field $B^{\mathrm{meas}}(t)$ for *M. infundibulum* giant axon (worm E) with $N_{\mathrm{avg}} = 1200$, following continuous magnetic monitoring of this worm with full laser and MW power for the preceding 24 hours. Peak-to-peak value of $B^{\mathrm{meas}}(t)$ for worm E is not statistically different from values for worms D, F, G, or H with posterior stimulation, indicating little if any negative effects from NV-diamond magnetic sensing over long time periods.

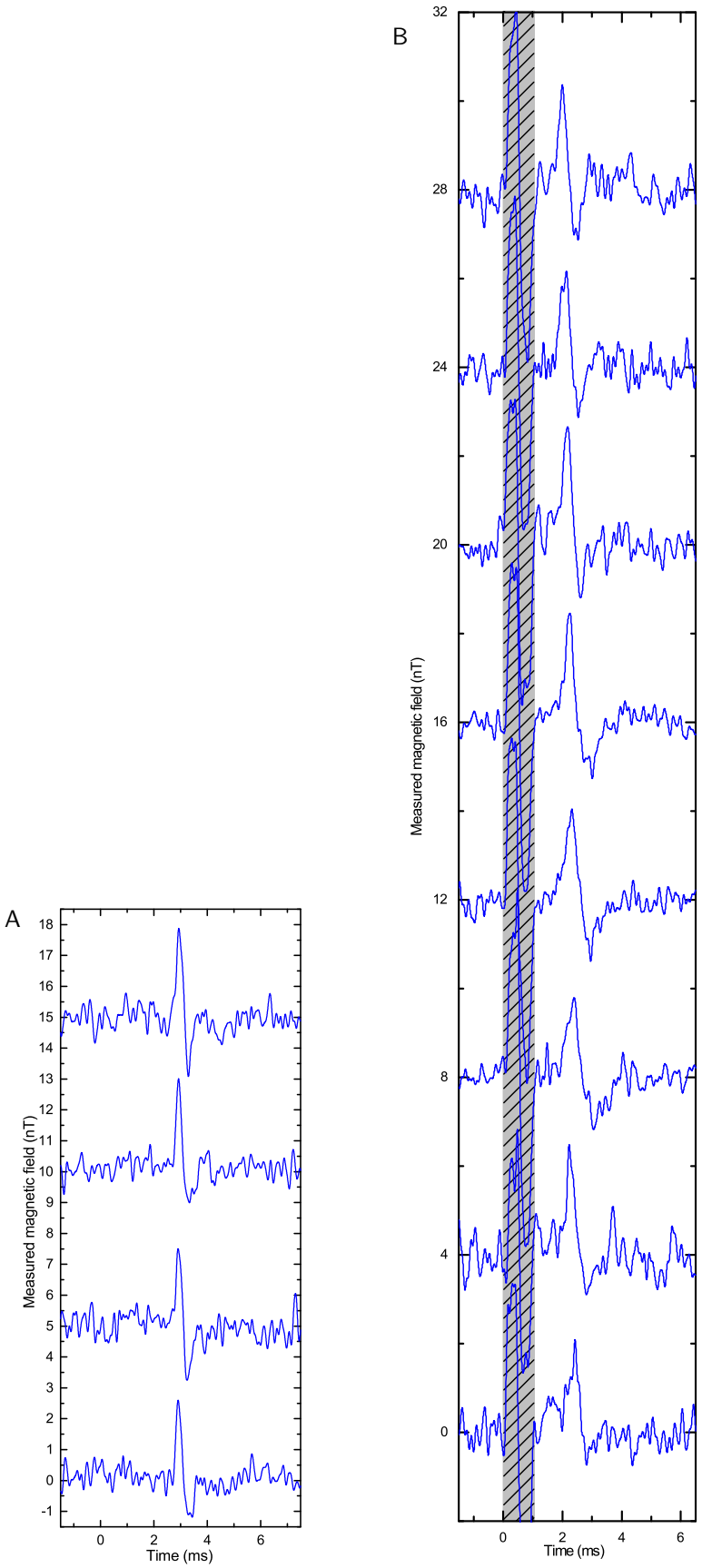

Fig. S11 | Technical replicates. (*A*) Four traces of $B^{\mathrm{meas}}(t)$ for worm B axon, each with $N_{\mathrm{avg}} = 150$, which are combined to make the data shown in Fig. 2*C*. Traces are offset for clarity with 5 nT spacing. (*B*) 8 additional technical replicates of the data in Fig. 2*D*. Traces are offset for clarity with 4 nT spacing.

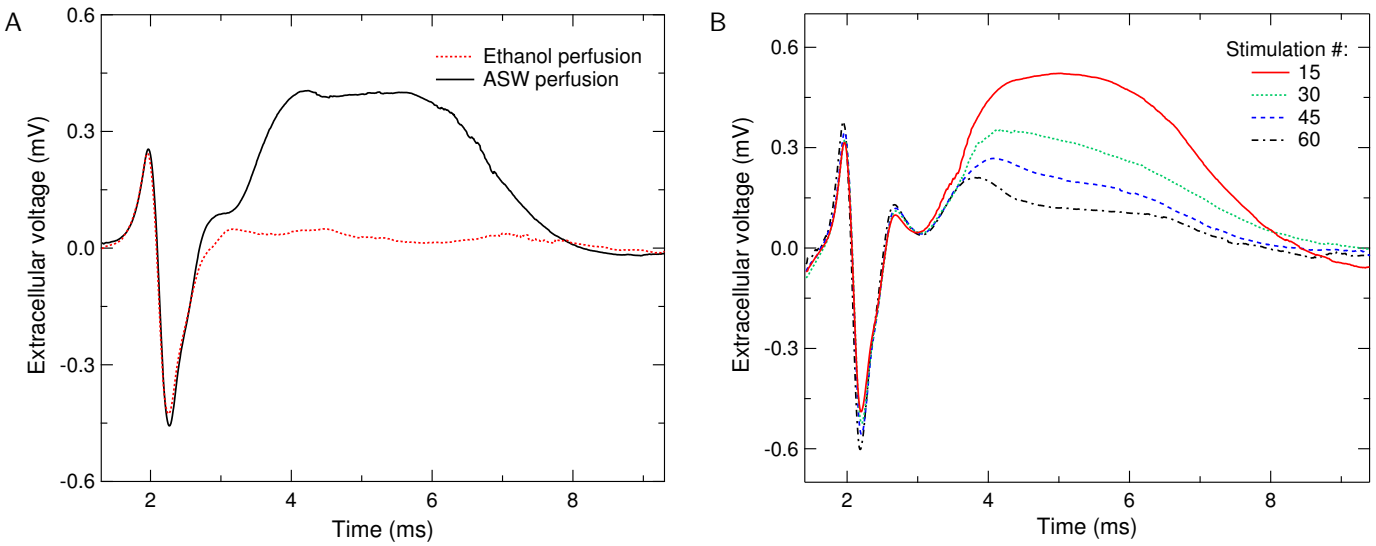

Fig. S12 | Muscle action. (*A*) Measured time traces of extracellular voltage $\Phi_{ex}^{meas}(t)$ following AP stimulation of intact *M. infundibulum* at the organism posterior with 0.5 Hz repetition rate. Red (dotted) trace shows a single trial $\Phi_{ex}^{meas}(t)$ from a worm perfused in a solution containing 10% ethanol and 90% ASW, which temporarily anesthetizes the muscles without affecting the giant axon AP. Black (solid) trace shows a single trial $\Phi_{ex}^{meas}(t)$ from the same worm stimulated roughly 20 minutes after perfusion is returned to 100% ASW. The delayed signal from ~3 ms to ~7 ms is not present in the anesthetized worm but returns after the worm's muscles become responsive again, strongly suggesting that the delayed signal is associated with muscle activity. (*B*) Measured time traces of single trial extracellular voltage $\Phi_{ex}^{meas}(t)$ following AP stimulation of intact *M. infundibulum* with conditions the same as in (*A*), with ASW perfusion. Data shown are for the 15th (red, solid), 30th (green, dotted), 45th (blue, dashed), and 60th (black, dot-dashed) AP stimulation. The signal from ~3 ms to ~7 ms is diminished by repeated stimulation, consistent with observed fatigue of the worm's muscles.

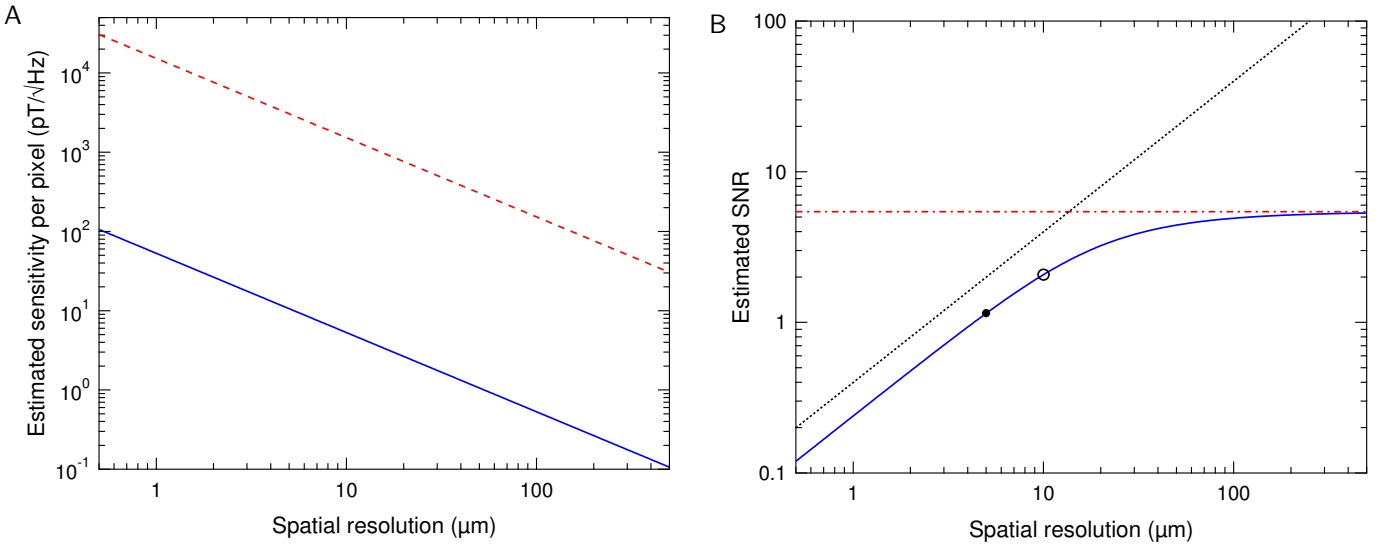

Fig. S13 | Estimated sensitivity and signal-to-noise ratio versus spatial resolution. (*A*) Scaling with spatial resolution l of the expected sensitivity of a magnetometer with volume-normalized sensitivity 34 nT $\mu m^{3/2} Hz^{-1/2}$ of the present instrument (red dashed line) and of a next-generation NV-diamond magnetic imager (blue solid line) with volume-normalized sensitivity 118 pT $\mu m^{3/2} Hz^{-1/2}$. Here l is defined in terms of the volume v of NV-diamond directly below the axon from which the NV fluorescence is imaged onto a single pixel, $v = l^2 d$, for an NV layer thickness $d = 5$ μm. (*B*) Magnetic signal-to-noise ratio (SNR) for a single pixel of a next-generation instrument with sensitivity 118 pT $\mu m^{3/2} Hz^{-1/2}$ versus spatial resolution l for NV layer thickness $d = 5$ μm and area $l \times l$. Blue trace shows scaling of estimated SNR from an axon of radius $r_a = 5$ μm with an estimated magnetic field of 1 nT at the axon surface, using Eqn. 19. Red dot-dashed trace denotes the asymptotic limit of the expected SNR arising from the inverse scaling of the magnetic field with distance ρ from the source: $B \propto 1/\rho$. Black dashed trace shows the expected SNR of a uniform 1 nT field versus l for $d = 5$ μm. Solid black dot marks expected SNR at $l = d = r_a = 5$ μm, and open circle marks expected SNR at $l = 10$ μm, the approximate crossover point between the two limiting regimes for $d = r_a = 5$ μm. The SNR of the same system using an identical magnetometer but with volume normalized sensitivity 34 nT $\mu m^{3/2} Hz^{-1/2}$ would provide signals with SNR reduced by a factor of 288.

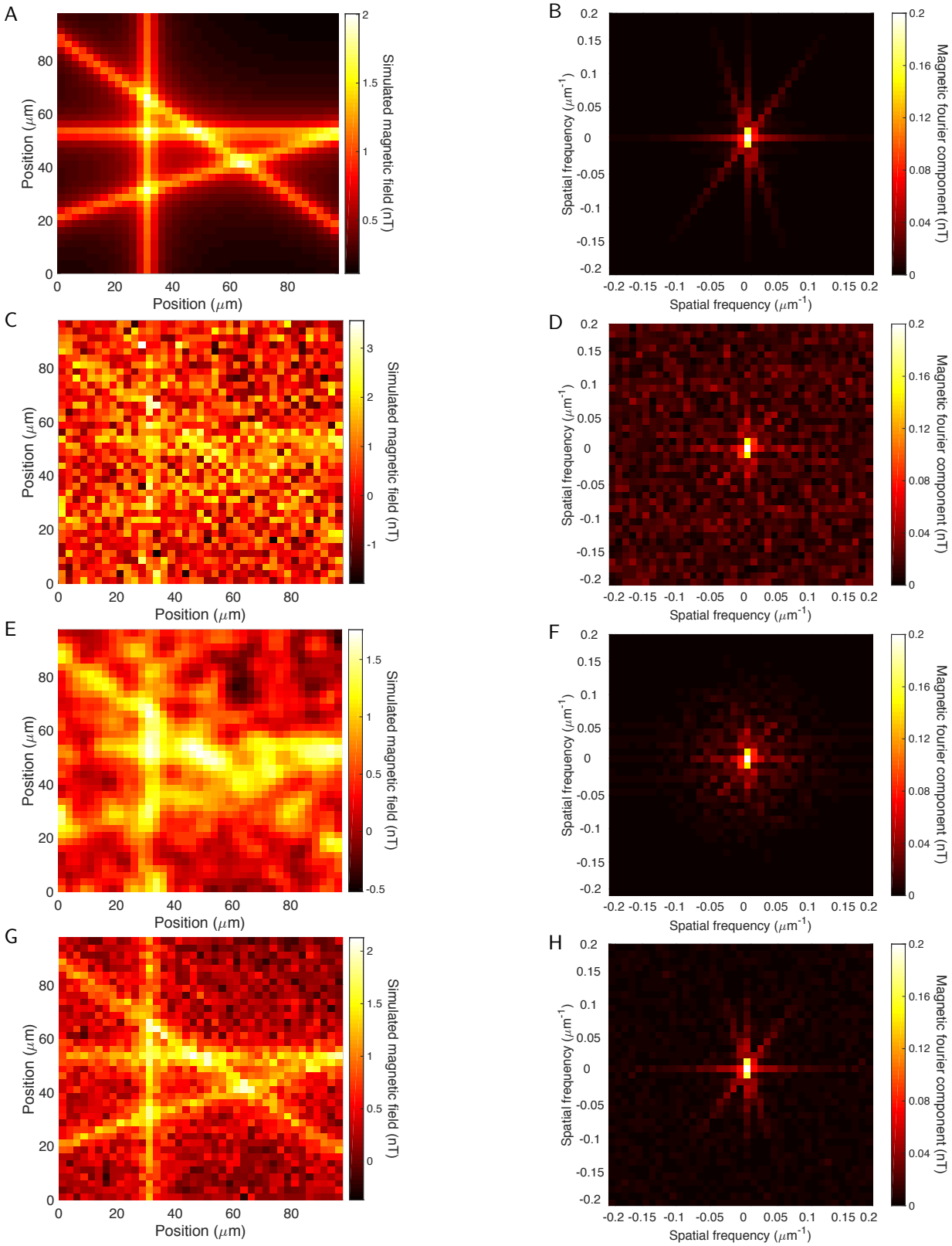

Fig. S14 | Simulated magnetic field maps from multiple neurons. (*A*) Calculated map of average magnetic field magnitude over NV layer from four wire-like axons with $r_a = 5$ μm firing near synchronously and each producing a maximum magnetic field amplitude of 1 nT at the axon surface. Each axon is placed directly on the NV sensing layer of thickness $d = 5$ μm. Field of view is 100 μm × 100 μm and pixel area is $l \times l = 2.5$ μm × 2.5 μm. (*B*) Spatial Fourier transform of field image in (*A*). The signal information is concentrated mainly at low spatial frequencies. (*C*) Simulated raw image of magnetic field map from (*A*) in the presence of Gaussian noise of 1.2 nT RMS deviation per pixel, as expected from a next-generation shot-noise-limited NV-diamond magnetic field imager with sensitivity 118 pT $\mu m^{3/2} Hz^{-1/2}$ and sampling rate 3 kS/s. A magnetic imager with volume-normalized sensitivity 34 nT $\mu m^{3/2} Hz^{-1/2}$ of the present instrument would require $N_{avg} = 83000$ averaged trials to yield an image with the same SNR. (*D*) Spatial Fourier transform of field map in (*C*). (*E*) Magnetic field map from (*C*) after processing with a spatial low-pass filter applied via convolution with a Gaussian kernel of 1-pixel (2.5 μm) standard deviation. (*F*) Spatial Fourier transform of field map in (*E*), displaying reduced contribution from noise at high spatial frequencies and preserved low-frequency signal information. (*G*) Magnetic field map from (*A*) simulated in the presence of $1/\sqrt{10}$ the noise in (*C*), demonstrating the SNR enhancement from averaging together $N_{avg} = 10$ trials using a next-generation magnetometer. (*H*) Spatial Fourier transform of field map in (*G*). To enhance higher spatial frequencies, the displayed color scale in (*B*), (*D*), (*F*), and (*H*) saturates the center pixel, which has value ~0.53 nT in all four images.

Reversal	**Result**	**Systematic ruled out**
Axon firing $\rightarrow$ Axon not firing	$B^{\text{meas}}(t) \rightarrow 0$	Any non-AP signal, including stimulation artifacts
$\Phi_{\text{ex}}(z_1, \rho_1, t) \rightarrow \Phi_{\text{ex}}(z_2, \rho_2, t)$	$B^{\text{meas}}(t) \rightarrow B^{\text{meas}}(t)$	Cross-talk of $B^{\text{meas}}(t)$ with $\Phi_{\text{ex}}(z, \rho, t)$
$\text{sgn}[\frac{dV_{\text{LIA}}}{df}] = 1 \rightarrow \text{sgn}[\frac{dV_{\text{LIA}}}{df}] = -1$	$B^{\text{meas}}(t) \rightarrow -B^{\text{meas}}(t)$	Any non-magnetic artifact including cross-talk of $B^{\text{meas}}(t)$ with $\Phi_{\text{ex}}(z, \rho, t)$
$\phi_{\text{LIA}} \rightarrow \phi_{\text{LIA}} + 180^\circ$	$B^{\text{meas}}(t) \rightarrow -B^{\text{meas}}(t)$	Cross-talk of $B^{\text{meas}}(t)$ with $\Phi_{\text{ex}}(z, \rho, t)$
$B_0 \rightarrow -B_0$	$B^{\text{meas}}(t) \rightarrow -B^{\text{meas}}(t)$	Magnetic artifact from motional coupling to small $\frac{\partial B_0}{\partial x}$ across diamond

Table S1 | Systematic checks