



Spatial information in large-scale neural recordings

Thaddeus R. Cybulski^{1*†}, Joshua I. Glaser^{1*†}, Adam H. Marblestone^{2,3}, Bradley M. Zamft⁴, Edward S. Boyden^{5,6,7}, George M. Church^{2,3,4} and Konrad P. Kording^{1,8,9}

¹ Department of Physical Medicine and Rehabilitation, Rehabilitation Institute of Chicago, Northwestern University, Chicago, IL, USA

² Biophysics Program, Harvard University, Boston, MA, USA

³ Wyss Institute, Harvard University, Boston, MA, USA

⁴ Department of Genetics, Harvard Medical School, Harvard University, Boston, MA, USA

⁵ Media Lab, Massachusetts Institute of Technology, Cambridge, MA, USA

⁶ Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA

⁷ McGovern Institute, Massachusetts Institute of Technology, Cambridge, MA, USA

⁸ Department of Physiology, Northwestern University, Chicago, IL, USA

⁹ Department of Applied Mathematics, Northwestern University, Chicago, IL, USA

Edited by:

Mayank R. Mehta, University of California, Los Angeles, USA

Reviewed by:

Tomoki Fukai, RIKEN Brain Science Institute, Japan

Sen Song, Tsinghua University, China

*Correspondence:

Thaddeus R. Cybulski and Joshua I. Glaser, Rehabilitation Institute of Chicago, Northwestern University, 345 E Superior St., Attn: Kording Lab Rm 1479, Chicago, IL 60611, USA

e-mail: cyb@northwestern.edu;

j-glaser@u.northwestern.edu

[†] These authors have contributed equally to this work.

To record from a given neuron, a recording technology must be able to separate the activity of that neuron from the activity of its neighbors. Here, we develop a Fisher information based framework to determine the conditions under which this is feasible for a given technology. This framework combines measurable point spread functions with measurable noise distributions to produce theoretical bounds on the precision with which a recording technology can localize neural activities. If there is sufficient information to uniquely localize neural activities, then a technology will, from an information theoretic perspective, be able to record from these neurons. We (1) describe this framework, and (2) demonstrate its application in model experiments. This method generalizes to many recording devices that resolve objects in space and should be useful in the design of next-generation scalable neural recording systems.

Keywords: neural recording, fisher information, resolution, technology design, optics, extracellular recording, electrical recording, statistics

1. INTRODUCTION

A concerted effort is underway to develop technologies for recording simultaneously from a large fraction of neurons in a brain (Alivisatos et al., 2013; Marblestone et al., 2013). For a technology to reach the goal of large-scale recording, it must gather sufficient information from each neuron to determine its activity. This suggests that neural recording methodologies should be evaluated and compared on information theoretic grounds. Still, no widely applicable framework has been presented that would quantify the amount of information large-scale neural recording architectures are able to capture. Such a framework promises to be useful when we want to compare the prospects of new recording technologies.

A neural recording technology can be judged by its ability to isolate signals from individual neurons. One common method of differentiating between signals from different neurons is through the neurons' locations: if the recording technique can determine that the signal sources are sufficiently far apart (by signal amplitude or other methods), then the signals likely come from different neurons. One can quantify this ability to spatially differentiate neurons using Fisher information, which measures how much information a random variable (e.g., a signal on a detector) contains about a parameter of interest (e.g., where the signal originated). Fisher information can be used to determine the optimal precision with which the parameter of interest

(the neural location) can be estimated¹. By calculating the Fisher information a technology carries about sources it records, one can determine how precisely neural locations can be estimated using this technology, and thus whether the neural activities can be distinguished in space.

Determining the Fisher information content of a sensing system allows determining the informatic limits of a technology in a given situation. These informatic limits, in turn, can guide technology design. For example, by quantifying the information content of an electrode array as a function of the spacing between electrodes, one could determine the spacing necessary to distinguish neural activities. Similarly, one can compare the information content of several optical recording approaches to determine the optimal technology for a given experiment.

Here we develop a Fisher information-based framework that characterizes neural recording technologies based on their abilities to distinguish activities from multiple neurons. We apply this framework to models of neural recording techniques, describe how the Fisher information scales with respect to recording geometries and other parameters, and demonstrate how this framework could be utilized to optimize experimental design.

¹ Fisher information is a theoretical calculation that determines the best a technology can do—signal separation techniques (e.g., Mukamel et al., 2009) are generally required to approach this optimum.

We demonstrate the utility of a Fisher information-based evaluation of neural recording technologies, which may inform the design and development of next-generation recording techniques.

2. FRAMEWORK

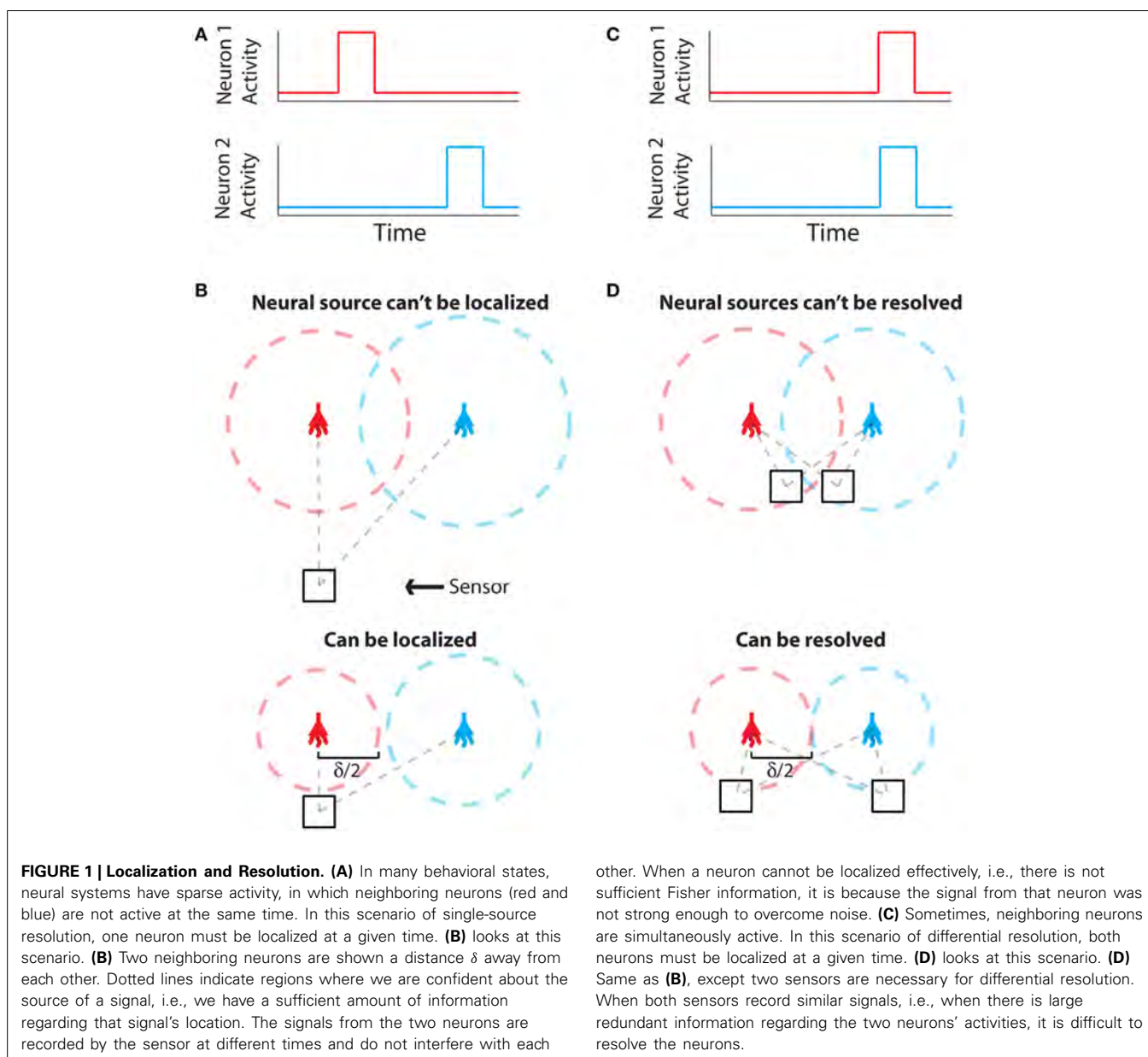
2.1. LOCALIZATION AND RESOLUTION

A fundamental concern in neural recording is localization, the ability to accurately estimate the location of origin of neural activity. Localization is a primary method of determining the identity of an active neuron.

The problem of establishing neural locations can be split into two separate regimes. One regime is when an active neuron has no active neighbors (**Figure 1A**). In this state, we are chiefly concerned with the ability to attribute the signal to the correct neuron (single-source resolution, Den Dekker and Van den Bos,

1997). This can be done by accurately localizing one activity at a given time on a background of noise (**Figure 1B**). The other regime is when two neighboring neurons are simultaneously active (**Figure 1C**). In this state, we are chiefly concerned with the ability to differentiate the two neurons, i.e., are there two clearly distinguished or one blurred neuron (differential resolution, Den Dekker and Van den Bos, 1997). This can be done by simultaneously localizing the activities of both neurons accurately (**Figure 1D**)².

²While we have been discussing differentiating neurons, the framework itself differentiates between point sources. In this paper, we make the assumption that separate point sources belong to separate neurons. In reality, it is possible that there could be separate signals from the cell body and dendrites that are perceived as different sources. These can be united using additional information (e.g., anatomical imaging or simultaneous activity).



Fisher information can be used to determine whether both scenarios are theoretically possible for a given technology. Here we treat both of these scenarios: first by calculating the Fisher information a sensing apparatus has about the location of a single neuron, and then expanding this framework to treat location parameters of multiple neurons. We address localization and resolution in the theoretical limit where the point spread function (PSF) is known, in order to study the limiting effects of neuronal and sensor noise on localization precision³.

Regardless of the number of neurons and sensors we are treating, Fisher information gives us a metric with which to evaluate a recording technology. Spatial information, the amount of information regarding the location of a source (i.e., a quantitative measure of localization ability), can be used to determine whether it is possible to correctly attribute an activity to its source (or multiple activities to multiple sources). In order to know the identity of a source, we must be confident about the location of origin of the activity with a positional error less than $\delta/2$, where δ is the distance from one neuron to another (Figures 1B,D). In terms of Fisher information, if we have sufficient information to locate the source of activity with a precision $\delta/2$, we can assign that activity to a single neuron that occupies that location.

2.2. FISHER INFORMATION: GENERAL PRINCIPLES

Fisher information is a metric that measures the information a random variable has about a parameter, and can be used to determine how well that parameter can be estimated. More precisely, Fisher information, $\mathcal{I}(\theta)$ is a measure of the information a random variable X , with distribution $f(X; \theta)$ parameterized by θ , contains about the parameter θ (Kullback, 1997):

$$\begin{aligned} \mathcal{I}(\theta) &= \mathbb{E} \left[\left(\frac{\partial}{\partial \theta} \log f(X; \theta) \right)^2 \middle| \theta \right] \\ &= \int \left(\frac{\partial}{\partial \theta} \log f(X; \theta) \right)^2 f(X; \theta) dx \end{aligned} \quad (1)$$

Intuitively, the more X changes for a given change in θ , the more information you will know about θ by observing X .

More generally, the Fisher Information a random variable X has about a parameter vector θ with k elements $[\theta_1 \dots \theta_k]$ can be represented by a $k \times k$ matrix with elements:

$$(\mathcal{I}(\theta))_{ij} = \mathbb{E} \left[\left(\frac{\partial}{\partial \theta_i} \log f(X; \theta) \right) \left(\frac{\partial}{\partial \theta_j} \log f(X; \theta) \right) \middle| \theta \right] \quad (2)$$

The elements of this matrix represent the information contained in a sample about a pair of parameters.

2.3. CRAMER-RAO BOUNDS

The optimal precision with which the parameter, θ , can be estimated is inversely related to the Fisher information contained about that parameter. More precisely, the variance of an unbiased estimator of a parameter is lower bounded by the Cramer-Rao bound (CRB) (Cramér, 1946):

$$\text{Var} \left[\hat{\theta}_i \right] \geq [\mathcal{I}(\theta)^{-1}]_{ii} \quad (3)$$

An important implication of this is that the CRB on θ_i not only depends on the information X contains about θ_i , but how similar θ_i 's effect on X is to the rest of the elements of θ . An off-diagonal term $(\mathcal{I}(\theta))_{ij}$ with large magnitude means that the parameters θ_i and θ_j are strongly correlated (or anti-correlated) in terms of their input on X . This will increase the CRB on estimating parameters θ_i and θ_j .

2.4. INDEPENDENCE AND SUMMATION

If two observations X_1 and X_2 are independently affected by θ , then the two Fisher information matrices about θ can be summed, as could be expected by the implications of independence on sample variance. This property allows us to easily apply our framework to situations with multiple samples, either by multiple sensors or multiple time points.

In the following sections, we will apply the above properties of Fisher information and CRBs to develop a framework for determining how precisely the location of neural activities can be estimated, and thus whether they can be distinguished. Note that, while we will describe the ability to distinguish neurons solely using spatial information, additional sources of information can be used, e.g., temporal information in optical (Pnevmatikakis et al., 2013) and electrical recordings (Lewicki, 1998) (see *Framework Discussion*).

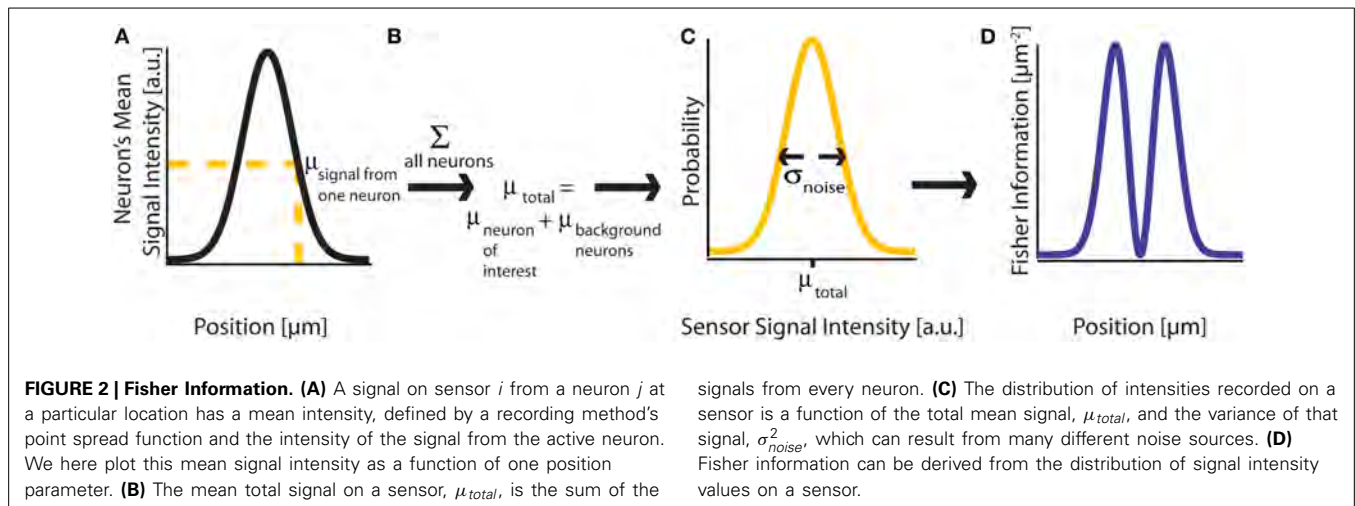
2.5. FISHER INFORMATION: SINGLE-SOURCE RESOLUTION

We first examine the situation where a single active source of some known intensity must be localized using an ensemble of sensors⁴. Here we observe a random variable, X , the value recorded at some sensor (e.g., in Volts). $f(X; \theta)$ then is the distribution of sensor values from repeated recordings of a neuron parameterized by θ . θ is a vector representing spatial (and other, e.g., intensity) parameters that characterize the neural signal. This resulting distribution $f(X; \theta)$ reflects both intrinsic variance of a neural signal as well as extrinsic factors such as other neurons and noise.

Here, Fisher information, $\mathcal{I}(\theta)$, measures how much the distribution of recorded sensor values $f(X; \theta)$ tells us about

³There exists a family of deconvolution techniques that estimate the PSF and use it to obtain a more accurate representation of the original signal (e.g., Colak et al., 1997; Onodera et al., 1998; Yan and Zeng, 2008; Broxton et al., 2013). In theory, with sufficient samples and knowledge of the PSF, one could obtain a perfect representation of a sparse signal in the absence of noise. This is not the case in practice, as signals are not only modified reversibly by PSFs, but are modified irreversibly by noise on neurons and detectors (e.g., Shahram and Milanfar, 2004; Shahram, 2005). In the presence of noise and other aberrations, it thus becomes difficult to isolate individual sources using deconvolution techniques, even when the PSF is known. Thus, it is interesting to determine the isolated effects of noise on recording methods. Moreover, as this Fisher information framework gives optimal bounds on precision with a known PSF, it can be used to determine how close to optimal a deconvolution algorithm performs.

⁴Activity in neural systems is often sparse (Bair et al., 2001; Cohen et al., 2010; Cohen and Kohn, 2011; Barth and Poulet, 2012; Denman and Contreras, 2014); this simplified scenario may be a useful model of neural systems.



the location of a signal's origin (Figure 2D). Intuitively, if a change in the signal origin's location would cause a large change in the recorded signal, then there will be a large amount of information about the location. However, if a change in the origin of the neural signal does not affect the recorded signal, there will be little information about the location of the neuron.

The CRB for a given parameter θ_i will tell us how precisely that location parameter can be estimated from the signal intensity. Assuming an unbiased estimator (the average estimate will be the true location), the best possible variance of the estimate is $[\mathcal{I}(\boldsymbol{\theta})^{-1}]_{ii}$. If we want to be confident that the estimated location of a given neuron's activity is within $\delta/2$ of its true location, as in Figure 1, the CRB on the estimate of distance must be less than $(\delta/4)^2$.⁵

Without assuming any prior knowledge, at least k variables are required to estimate k parameters, as the system is underconstrained with smaller numbers of samples. In our case, we need multiple sensors in order to estimate a neuron's location. If the sensors have independent noise—an assumption we use in our demonstrations—the information matrices can be summed (See *Independence and Summation*).

2.6. FISHER INFORMATION: DIFFERENTIAL RESOLUTION

In the scenario of multiple neurons acting simultaneously, we are interested in using signals recorded from an ensemble of sensors to estimate the location parameters of each neuron. That is, $\boldsymbol{\theta}$ now represents the location parameters of all neurons in the system, and $f(X; \boldsymbol{\theta})$ represents the distribution of signal intensities on a sensor given all of the neurons in the system. We can then construct a Fisher information matrix to determine the precision with which each parameter can be estimated. If each sensor recording is affected by n neurons, each with k parameters, the Fisher information matrix will be $nk \times nk$. The CRB calculated in this scenario will be most applicable to determining whether technologies are able to effectively record from a population of neurons.

⁵95% confidence under Gaussian assumptions.

2.7. POINT SPREAD FUNCTIONS AND SIGNAL INTENSITY DISTRIBUTIONS

To determine the spatial Fisher information, we must know the distribution of signals on a sensor given the location of the activity, $f(X; \boldsymbol{\theta})$. In this section, we derive the general form of $f(X; \boldsymbol{\theta})$ based on the PSF of a technology.

The signal measured by many recording systems is well-approximated as a linear function of the signals from each neuron in a population (Johnston et al., 1995; Cremer and Masters, 2013), i.e., the total sensor signal is the sum of the individual neural signals weighted by the magnitude of their individual effects on the sensor (Figures 2A,B). We thus only consider linear interactions; it should be noted that the Fisher information framework is also compatible with non-linear interactions (e.g., sensor saturation). For N neurons and M sensors in a system, in the absence of noise, the signal on any particular sensor can therefore be described as:

$$\mathbf{x} = \mathbf{W}\mathbf{a} + \boldsymbol{\epsilon} \quad (4a)$$

where \mathbf{x} is the vector of signals on sensors $[X_1, \dots, X_M]$, $\boldsymbol{\epsilon}$ is the vector of noise on each sensor $[\epsilon_1, \dots, \epsilon_M]$, which arises from neural and sensor noise, and \mathbf{a} is the vector of signals from neural activities, $[I_1, \dots, I_N]^T$, e.g., the fluorescent signal produced due to neural activity in optical techniques or the voltage signal in electrical techniques. \mathbf{W} is the matrix of PSFs:

$$\mathbf{W} = \begin{bmatrix} w(\mathbf{d}^{1,1}) & \dots & w(\mathbf{d}^{1,N}) \\ \vdots & \ddots & \vdots \\ w(\mathbf{d}^{M,1}) & \dots & w(\mathbf{d}^{M,N}) \end{bmatrix} \quad (4b)$$

where w is the PSF, which depends on the location of the neuron relative to the sensor and other parameters of a recording modality (e.g., light scattering). $\mathbf{d}^{i,j}$ is a vector that gives the location of neuron j relative to sensor i . It has elements $[\mathbf{d}_1^{i,j} \dots]$ that describe single location parameters of $\mathbf{d}^{i,j}$.

Combing Equations 4a and 4b, we can write the total signal on a sensor i as

$$X_i = \sum_j I_j w(\mathbf{d}^{i,j}) + \epsilon_i \quad (5a)$$

We can write a function $f(X_i)$ that characterizes the distribution of signal intensities on a sensor. Here, we assume that the noise, ϵ_i , can be approximated by a zero-mean Gaussian with variance σ_{noise}^2 , so that:

$$f(X_i; \theta) = \mathcal{N} \left(\sum_j I_j w(\mathbf{d}^{i,j}), \sigma_{noise}^2 \right) \quad (5b)$$

where $\mathcal{N}(\mu, \sigma^2)$ signifies a normal distribution (**Figure 2C**). θ is the vector of parameters that we are estimating. It can include any I_j and any elements of any $\mathbf{d}^{i,j}$. This allows us to calculate the Fisher information in signal X_i about location (or intensity) parameters of neurons using Equations 1 or 2 (**Figure 2D**). Note that the Gaussian noise assumption allows for simplifications in the Fisher information calculation (see *Supplementary Material* for derivation).

It is also important to note that, as long as they can be analytically described, all types of noise (of which there are many; see *Supplementary Material* for further discussion) can be incorporated into this framework. This flexibility in noise sources makes this framework especially relevant for neural recording.

3. FRAMEWORK DISCUSSION

Here we have described a framework to quantitatively approach the challenges of large-scale neural recording and determine the necessary experimental parameters for potential recording modalities. This framework extends previous work applying Fisher information to individual imaging techniques (e.g., Helstrom, 1969; Winick, 1986; Ober et al., 2004; Aguet et al., 2005; Shahram, 2005; Shahram and Milanfar, 2006; Marengo et al., 2009; Sanches et al., 2010; Mukamel and Schnitzer, 2012; Quirin et al., 2012; Shechtman et al., 2014). For example, many studies have used Fisher information to examine the theoretical optimal resolution of specific optical imaging techniques (Helstrom, 1969; Winick, 1986; Ober et al., 2004; Aguet et al., 2005; Shahram, 2005; Shahram and Milanfar, 2006; Marengo et al., 2009; Mukamel and Schnitzer, 2012; Quirin et al., 2012; Shechtman et al., 2014). However, using Fisher information to optimize other neural recording technologies, while occasionally done (e.g., MRI in Sanches et al., 2010), is not as common. Moreover, as many of the previous approaches are optics-centric, they generally do not consider the effects of recording in biological tissue, a central concern in neuroscience.

We expand on previous work by considering a PSF and noise model based on recording in neural tissue, and then using a Fisher information-based approach to establish signal separability. It is able to describe the information content of neural recording technologies that separate sources based on location, of which there are many. This information content can then be used to evaluate a technology's ability to separate sources. Such a framework

promises to be useful in evaluating and comparing novel and established recording technologies.

Given this framework's reliance on signal modulation by PSFs, it neglects other ways that sources can be separated, such as color (Hampel et al., 2011) or spike waveform. Some of this information could be made compatible with our framework via virtual recording channels, e.g., in time. While these types of non-spatial information are not considered here, they may be necessary to separate sources under certain recording situations, e.g., where the dendrites of one neuron produce a signal within the CRB of the cell body of another neuron. In an extreme case, proposed intracellular molecular recording devices have no spatial information, but could still effectively separate signals (Kording, 2011; Zador et al., 2012). While spatial Fisher information is an attractive method of evaluating neural recording techniques, it is important to remember these limitations when considering non-spatial techniques.

In addition, the CRBs described here only consider unbiased estimators. That is, they only provide a lower bound on localization ability when there are no prior assumptions about neurons' locations. It is possible to be more precise than the CRB if the estimator is biased (i.e., if assumptions are made about neurons' locations, or neurons' locations are constrained). There is work on Bayesian Cramer Rao Bounds (Van Trees, 2004; Dauwels, 2005) and bounds on parameter estimation with constraints (Gorman and Hero, 1990; Matson and Haji, 2006) that could be applied to better understand the capabilities of recording technologies.

This framework is particularly suited to the evaluation of novel techniques due to its general nature; it is applicable to any technique where a spatial PSF can be measured and the system's noise distribution can be either modeled or explicitly described. For instance, advanced optical techniques (Ahrens et al., 2013; Prevedel et al., 2014), ultrasound, and MRI have all been proposed as potential large-scale neural recording techniques (Marblestone et al., 2013; Seo et al., 2013). With a PSF describing how signals from different positions in the brain reach a sensor (some discussion in Jensen, 1991; Smith and Lange, 1998; Engelbrecht and Stelzer, 2006; Shin et al., 2009; Qin, 2012, and Prevedel et al., 2014) and further quantification of recording noise, this framework could easily be applied to determine bounds on signal separability for those techniques.

Ultimately, the utility of this approach is dependent on the quality of PSFs and noise models we have. For some techniques, these are well-described (especially PSFs); for others, these are poorly understood. As models of neural recording techniques advance, the predictions of this technique will become more accurate.

4. DEMONSTRATIONS

Here, we demonstrate the utility of the Fisher information framework for analysis of neural recording technologies. We provide demonstrations of the use of Fisher information in the cases of single-source and differential resolution. We first calculate the spatial Fisher information of a single source in simple recording setups for several model recording methods. We next demonstrate more realistic uses of the Fisher information

framework: optimal technology design, technology comparison, and estimating locations when the neural activity's intensity is unknown.

4.1. ASSUMPTIONS

For our demonstrations, we make several assumptions. First, we assume that all activity from the neuron of interest, including the noise, is part of the signal of interest. Thus, the total noise is a function of the sensor noise plus the noise of all neurons except for the neuron of interest. In order to create an accurate model of a neural recording technology, we must know how all sources of noise affect the recorded signal, and also the relation between the noise and the intensity of the neural activity. Because these are in general not known, we make further assumptions in our simulations.

In regards to neural activities, we assume that every active neuron has the same activity I_0 (except when otherwise stated), while non-active neurons have no activity, that the neuron of interest, k , is active at the moment we sample, and that other neurons are active at a uniform rate. We assume noise sources from neurons are independent, so that:

$$\sigma_{noise}^2 = \sum_{j \neq k} \sigma_j^2 \quad (6)$$

There are many sources of noise, both on neurons and sensors, that could be included; these are discussed in the *Supplementary Material*. For our demonstrations, we consider signal dependent noise that can arise from neurons and/or sensors. Specifically, for analytic simplicity, we only consider noise that has a standard deviation proportional to the mean signal: $\sigma_j^2 \propto I_0^2 (w(\mathbf{d}^{i,j}))^2$. We use these simplifying assumptions so that the magnitudes of the fluorescence (optical) and waveform voltage (electrical) have no influence on the final information theory calculations (and the relationship between these magnitudes and the noise is not in general well-understood). We emphasize that these simulation assumptions are implemented to simply demonstrate the use of this framework; more realistic outputs could be found using more complex, realistic noise models.

4.2. SINGLE NEURON LOCALIZATION

Here we calculate Fisher information of recording technologies using a single neuron and simple sensor arrangements as an illustration of our framework. We look at three technologies: (1) electrical recording, a traditional neural recording modality, (2) wide-field fluorescence microscopy, a traditional optical approach, and (3) two-photon microscopy, a modern optical approach. These examples are chosen for their relative simplicity and ability to illustrate the flexibility of a Fisher information approach to modeling neural recording.

For any technology, the aim is for there to be, across all sensors, sufficient information about every location in the brain in order to identify a neuron firing in that location. Thus, for an individual sensor, it can be better to have sufficient (enough to identify a neuron, as in **Figure 1**) information spread over a large area than excessive information about a small area. This suggests that experimental designs could be modified to get sufficient information for the required task. For example, an optical technology may have extra information at low depths, but insufficient information at large depths. In this case, the PSF could be modulated (e.g., Quirin et al., 2012) to decrease low-depth information (making those images blurrier), while increasing high-depth information.

4.3. ELECTRICAL SENSING

The electrical potential from an isolated firing neuron decays approximately exponentially with increasing distance (Gray et al., 1995; Segev et al., 2004), at least at short distances. Here, we model a simple electrical system: an isotropic electrode with spherical symmetry (**Figure 3A**). In this isotropic approximation, the PSF has an exponential decay with radial distance from the electrode tip (**Figure 3B**; PSF taken from **Table 1**, using parameters found in **Table 2** and **Figure 3**).

For electrical recording, estimators of location parameters have the lowest standard deviation σ_x and σ_y when in-between two electrodes, and the lowest σ_z when directly above or below an electrode (**Figures 3D,E**). Generally, we see that electrical recordings provide relatively weak information over a relatively wide area. In fact, we find that, in “worst-case” regions, standard electrode arrays should have difficulty localizing a source within the bounds required to discriminate between neighboring

Table 1 | Point spread functions of recording modalities.

Electrical	$w_{el}(r) = \exp\left(\frac{-r}{C_{el}}\right)$
Optical: Wide-field fluorescence microscopy	$w_{wff}(\ell, z) = \frac{Q}{2\pi} \exp\left(\frac{-z}{C_{op}}\right) \frac{1}{(s_{defocus}^2 + s_{dif}^2 + s_{scat}^2)} \times \exp\left(\frac{-\ell^2}{2(s_{defocus}^2 + s_{dif}^2 + s_{scat}^2)}\right)$ $s_{defocus} = \frac{D_{lens} \cdot (z_0 - z)}{2z_0}, s_{dif} = \frac{0.42\lambda \cdot z}{D_{lens}}, s_{scat} = \gamma z$
Optical: 2-photon microscopy	$w_{2P}(\ell, z) = \frac{1}{\pi} \frac{1}{(s_{defocus}^2 + s_{dif}^2 + s_{scat}^2)} \times \left(Q \exp\left(\frac{-z}{C_{op}}\right) \exp\left(\frac{-\ell^2}{2(s_{defocus}^2 + s_{dif}^2 + s_{scat}^2)}\right) \right)^2$

Analytic expressions are given for PSFs. r is the distance in any radial direction from the electrode, and ℓ is the lateral distance from the center of the lens for optical techniques. Note that $r^2 = x^2 + y^2 + z^2$ and $\ell^2 = x^2 + y^2$. C_{el} is the spatial constant of electrical decay. C_{op} is the spatial constant of optical decay. $s_{defocus}^2$, s_{dif}^2 , and s_{scat}^2 are the variance of the spread of optical light due to defocusing, diffraction, and scattering, respectively. D_{lens} is the diameter of a lens. λ is the wavelength of the light. z_0 is the focus depth, and Q is the light flux (area per photon).

Table 2 | Simulation Parameter Values.

Parameter	Value
C_{el}	28 μm (Gray et al., 1995; Segev et al., 2004)
D_{lens}	300 μm (within current dimensions)
λ (wide-field)	633 nm (visible light)
λ (2-photon)	800 nm (infrared light)
γ (wide-field)	0.15 (Orbach and Cohen, 1983; Tian et al., 2011)
C_{op} (wide-field)	100 μm (with 515 nm light) (Theer and Denk, 2006)
γ (2-photon)	0.002 (with 725 nm light) (Chaigneau et al., 2011)
C_{op} (2-photon)	200 μm (with 909 nm light) (Theer and Denk, 2006)

neurons. Given that current arrays generally require more information than a single sample of signal intensity to sort spikes (e.g., waveform shape is used), this is an expected result.

4.4. OPTICAL SENSING

4.4.1. General information

Optical recording of neural activity generally relies on fluorescent dyes that are sensitive to activity. In order to measure this signal, a neuron must be illuminated with light in the dye's excitation spectrum. Light is then emitted by the dye at a distinct, longer (lower energy) wavelength, which is picked up by a photodetector. Optical signal transmission is subject to absorption, scattering, and diffraction, which degrade the emitted signals with distance. Absorption of light effectively cause an exponential decrease in intensity of detected photons as light travels through a medium (Lambert and Anding, 1892; Theer and Denk, 2006). Scattering can affect light in multiple ways; high-angle scattering diverts photons from the detector and produces an effect similar to absorption, while low-angle scattering causes blurring of the image on the detector. This blurring increases approximately linearly with depth into the tissue (Tian et al., 2011). Finally, diffraction results when light passes through an aperture, creating the finite-width Airy disk (Airy, 1835). In our optical PSFs, we assume scattering and diffraction result in Gaussian blurring (Thomann et al., 2002; Tian et al., 2011). Our PSFs assume imaging through a single homogeneous medium; in practice, tissue inhomogeneity and refractive index mismatch can produce additional aberrations in the absorption, scattering, and diffraction domains that we do not model here.

In a typical optical setup, a lens focuses a set of photons from one point in space onto a corresponding point behind the lens. This phenomenon can be used either to focus incident light onto a desired location for illumination, or to focus emitted light from the focal plane onto a photodetector for imaging. Photons from outside the focal plane will be blurred, and this blurring increases linearly as distance from a focus point increases (Torreao and Fernandes, 2005; Kirshner et al., 2013). We also assume defocusing results in Gaussian blurring (Torreao and Fernandes, 2005; Kirshner et al., 2013).

4.4.2. Wide-field fluorescence microscopy

Neural activity in a focused optical system is generally sensed using fluorescent dyes, which require some excitatory light. In the canonical optical example of wide-field microscopy, an entire

volume is illuminated (Figure 4A). The PSF for this technology takes the above effects of absorption, scattering, diffraction, and defocusing into account; we assume total illumination so that the PSF here models the spread of the emission light (Figure 4B, PSF taken from Table 1 using parameters found in Table 2).

For optical recording with a simple lens, estimators of location parameters have lowest standard deviation σ_x , σ_y , and σ_z when centered above the imaging system in the focal plane (Figures 4D,E). For large depth, the ability to distinguish locations decreases rapidly due to photon loss caused by scattering and absorption (Figures 4D,E). For medium depth ranges, scattering blurs the image, even on the focal plane. These phenomena decrease the utility of deep focal-plane wide-field optics in tissue. At shallower focal depths, optical recordings provide a large amount of information on the focal plane, while carrying relatively little information about sources out of the focal plane (Figures 4D,E).

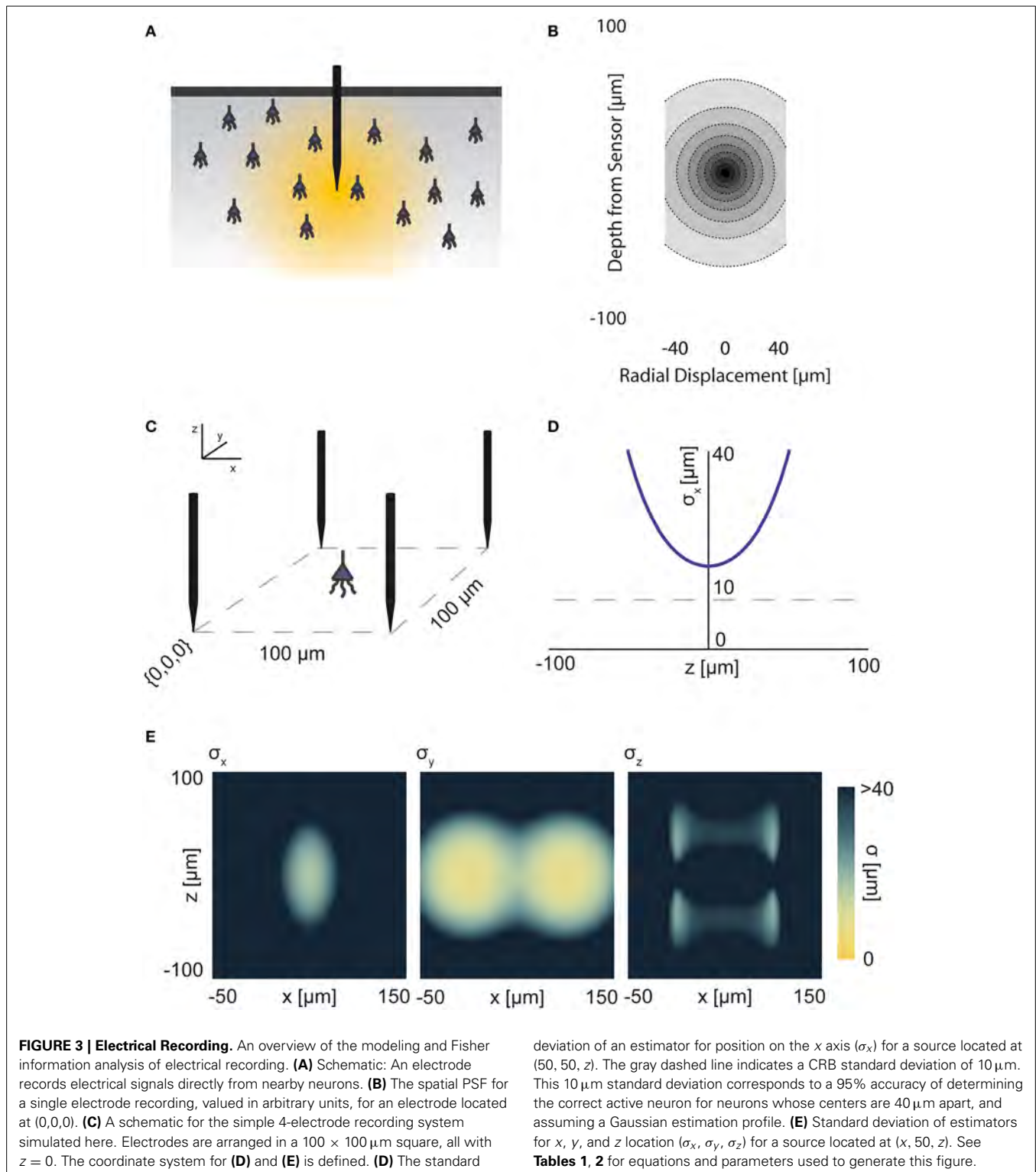
4.4.3. Two-photon microscopy

In two-photon microscopy, long-wavelength incident light (i.e., composed of low-energy photons) is focused onto a single point of interest to excite fluorophores in that area (Figure 5A). In order for the fluorophore to emit light, two low-energy photons must be absorbed nearly simultaneously; the likelihood of this event is proportional to the square of the intensity of incident light at a point. Effectively, this concentrates the area of sufficient illumination to a volume nearby the focal point of the incident beam (while increasing the illumination power requirements) (Helmchen and Denk, 2005). Like with wide-field fluorescence microscopy, the PSF is a function of defocusing, absorption, and scattering (Figure 5B, PSF taken from Table 1 using parameters found in Table 2). We assume total photon capture so that the PSF here models the spread of the excitation light.

For two-photon microscopy, estimators of location parameters have lowest standard deviation σ_x , σ_y , and σ_z just above and below the focal plane (Figures 5D,E). Perhaps counter-intuitively, there are extremely-high or undefined σ 's along the focal plane. This is due to our simplified recording setup (Figure 5C): given the tightly-focused PSF for two-photon microscopy, sources very close to the focal plane of our setup are effectively only "seen" by one sensor. Thus, we cannot gather meaningful information about the source's three location parameters, resulting in a singular or near-singular Fisher information matrix. In practice, this is alleviated by either decreasing the pitch of sensed regions or applying magnification to the sample, which we do not model here. We also see a reduced dependence on focal depth when compared to a wide-field imaging setting, as expected (Figure 5D).

4.5. TECHNOLOGICAL OPTIMIZATION

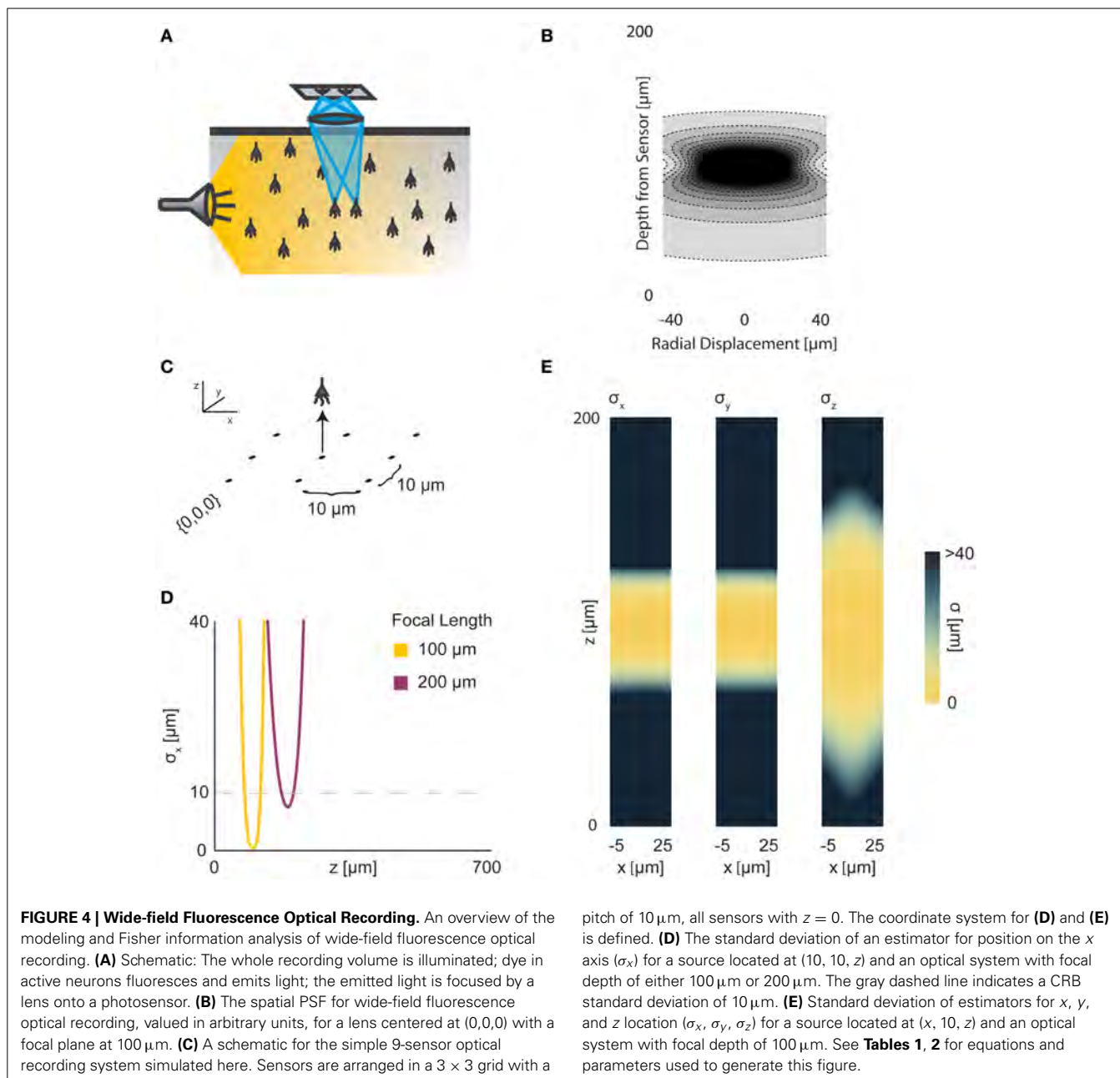
This example will demonstrate the ability to use Fisher information to ask questions about the necessary experimental parameters of neural recording technologies. In particular, we will use Fisher information to examine sensor placement in electrical recording. In order to successfully record activity from every neuron in a volume, we must place sensors so that they extract sufficient information about every neural location in that volume. That is, the CRB regarding the ability to estimate the location



of each point in a volume must be below some threshold for localization.

Here, we simulate several possible arrangements of electrical sensors and evaluate the information that these systems provide about different locations in a volume. Specifically, we look at

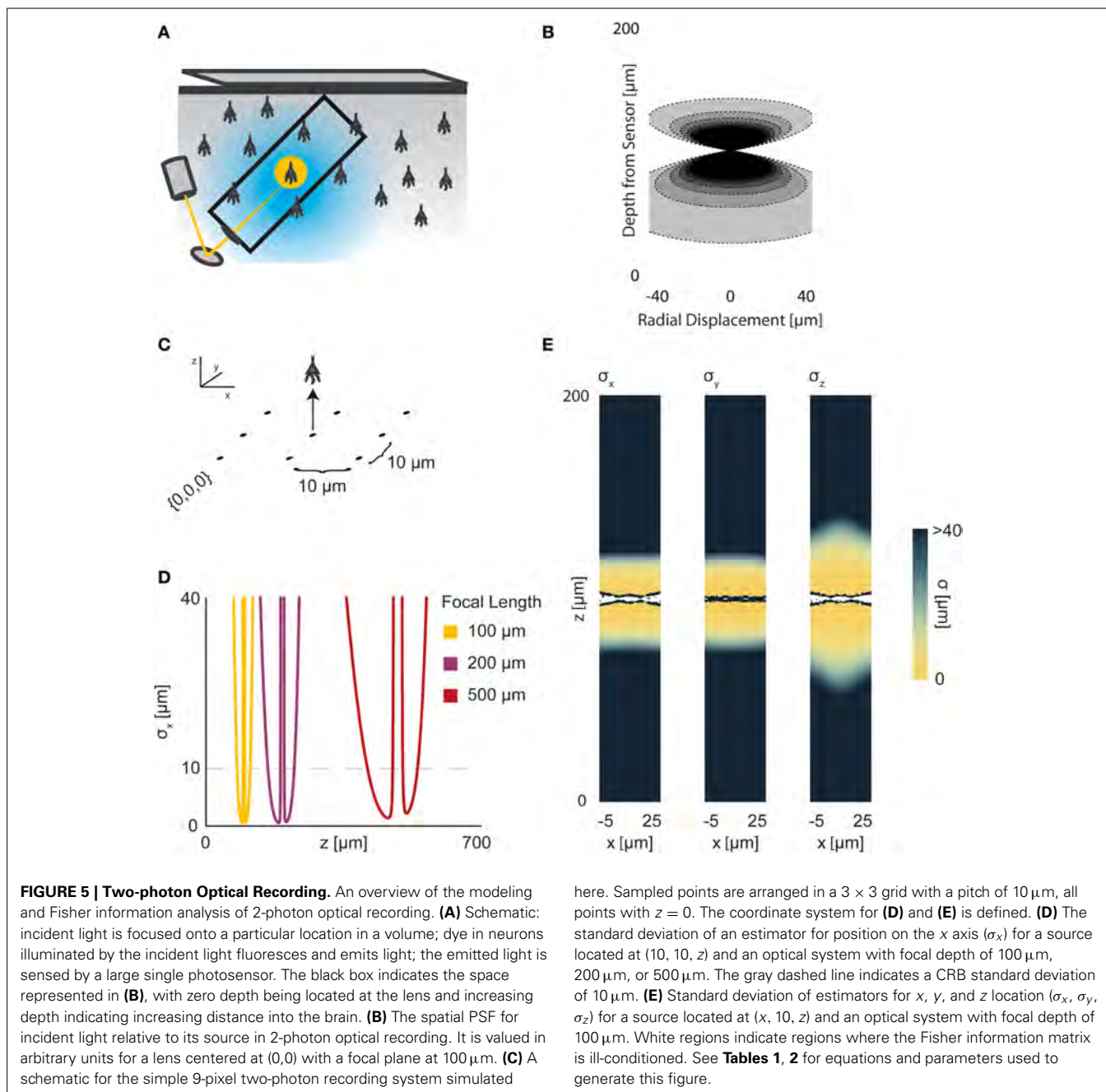
five electrode arrangements: (1) electrodes evenly distributed in an equilateral grid (Grid electrodes); (2) randomly placed electrodes (Random electrodes); (3) electrodes evenly distributed in a plane (Planar electrodes); and (4 and 5) two arrangements of columns of electrodes, where electrodes are densely packed within



a column, and these columns are arranged in a grid (Zorzos et al., 2012) (Column electrodes) (Figure 6A). Here, we assume that noise is independent between sensors, i.e., noise is all on the sensor. Under this assumption, each electrode takes an independent sample of a signal; information about the location of the source of that signal is then additive across sensors. Fisher information here is thus the information the entire ensemble of electrodes provides about a point. In this simplified example, we determine localization, rather than resolution, capabilities, which corresponds to the common situation of sparse neural firing. Multiple sources would necessarily reduce the amount of information contained about individual sources and would be geometry-dependent.

In this simplified simulation, Grid electrodes and Random electrodes have the best performance, as they sample space

uniformly (Grid) or almost uniformly (Random) (Figure 6B). Due to the regular nature of Grid electrodes, there is the added benefit of a guaranteed lower bound for information carried about locations in a volume. Planar electrodes are able to estimate a small fraction of locations very well, but carry very little information about most locations in a volume. Columnar electrodes, in general, have the interesting property that the z coordinate can be estimated more accurately, due to the density of electrodes in this direction. It's also important to note that the feasibility of Columnar electrodes will likely depend on the spacing between shanks. As the shanks move closer together (e.g., the bottom row compared to the fourth row), a greater number of neurons will be able to be distinguished. The use of this Fisher information framework promises to inform sensor placement decisions.

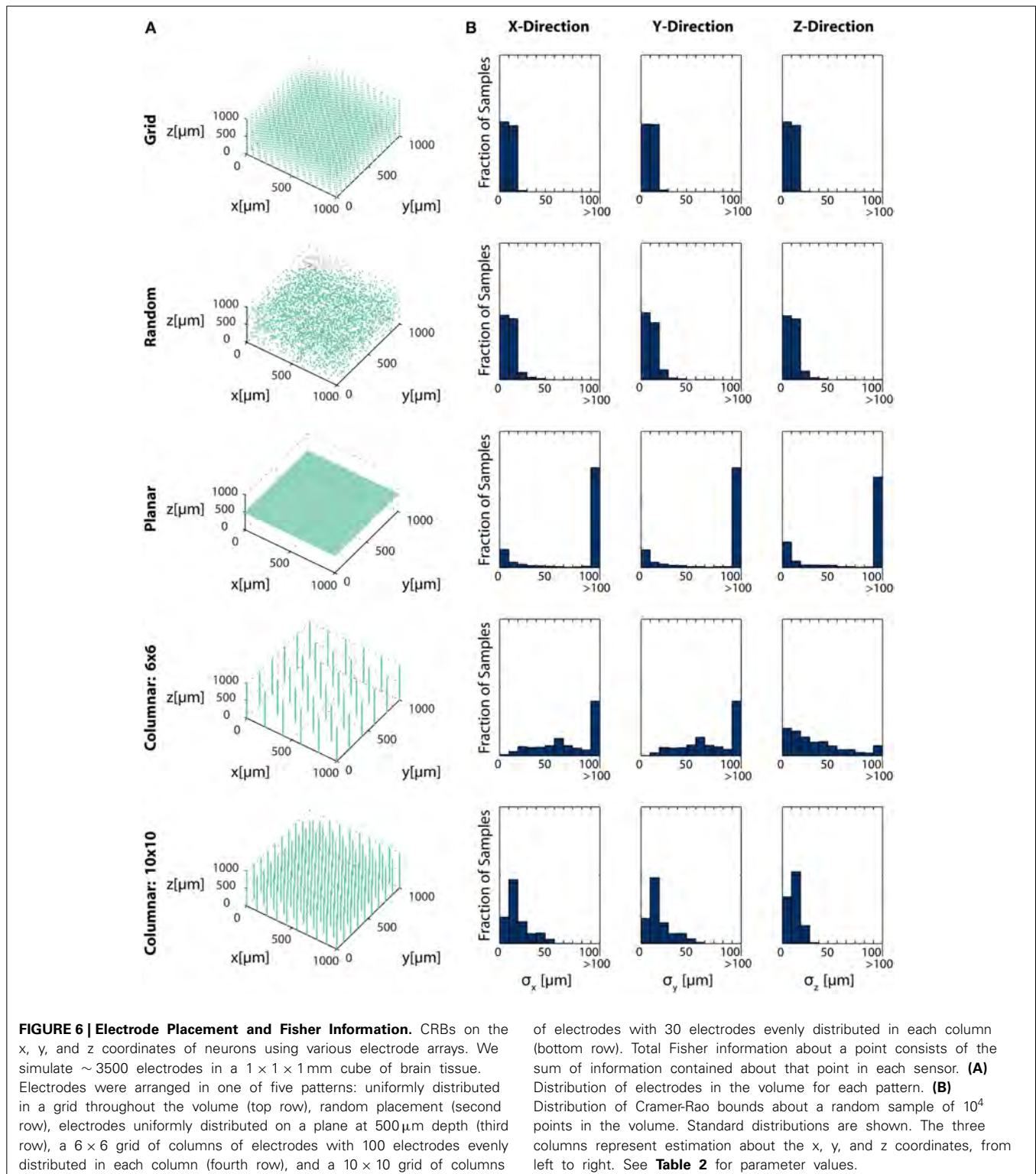


4.6. TECHNOLOGY COMPARISON

In this example, we demonstrate the use of Fisher information for determining resolution ability rather than localization ability. This example will demonstrate the ability to use Fisher information to compare technologies. In order to determine appropriate technologies for a given situation, it is necessary to know which technology will maximize the information output, and where information will be concentrated for a given technology.

Here we apply this Fisher information framework to a two-source, multi-sensor setup for both wide-field fluorescence and two-photon microscopy in order to determine performance over

depth (**Figure 7**). We find, perhaps confirming intuition, that wide-field and two-photon fluorescence perform similarly for shallow sections, but performance of wide-field fluorescence microscopy degrades significantly at a depth of 500 μm while two-photon performs well at this depth. Interestingly, both methods contain a large amount of information not only about signals near the focal point, but also about sources nearby the lens. This implies that signals could be recovered from out-of-focus samples given proper recording conditions. While this demonstration yielded the expected results, this framework could be used to compare existing technologies in novel situations, or to compare novel technologies.



4.7. ESTIMATION WITH UNCERTAIN SIGNAL INTENSITY

In previous sections, for the sake of simplicity, we have assumed a known, constant I_0 representing the intensity of any active source in the field. Here, we demonstrate the use of our framework without this assumption, using Fisher information to characterize a

sensing system's ability to localize a source with an uncertain intensity. To do this, we must determine the CRBs on estimators of 4 parameters: the three Cartesian coordinates of a source, along with the source's intensity, i.e., θ is a 4-element vector. We provide a simple demonstration of this technique

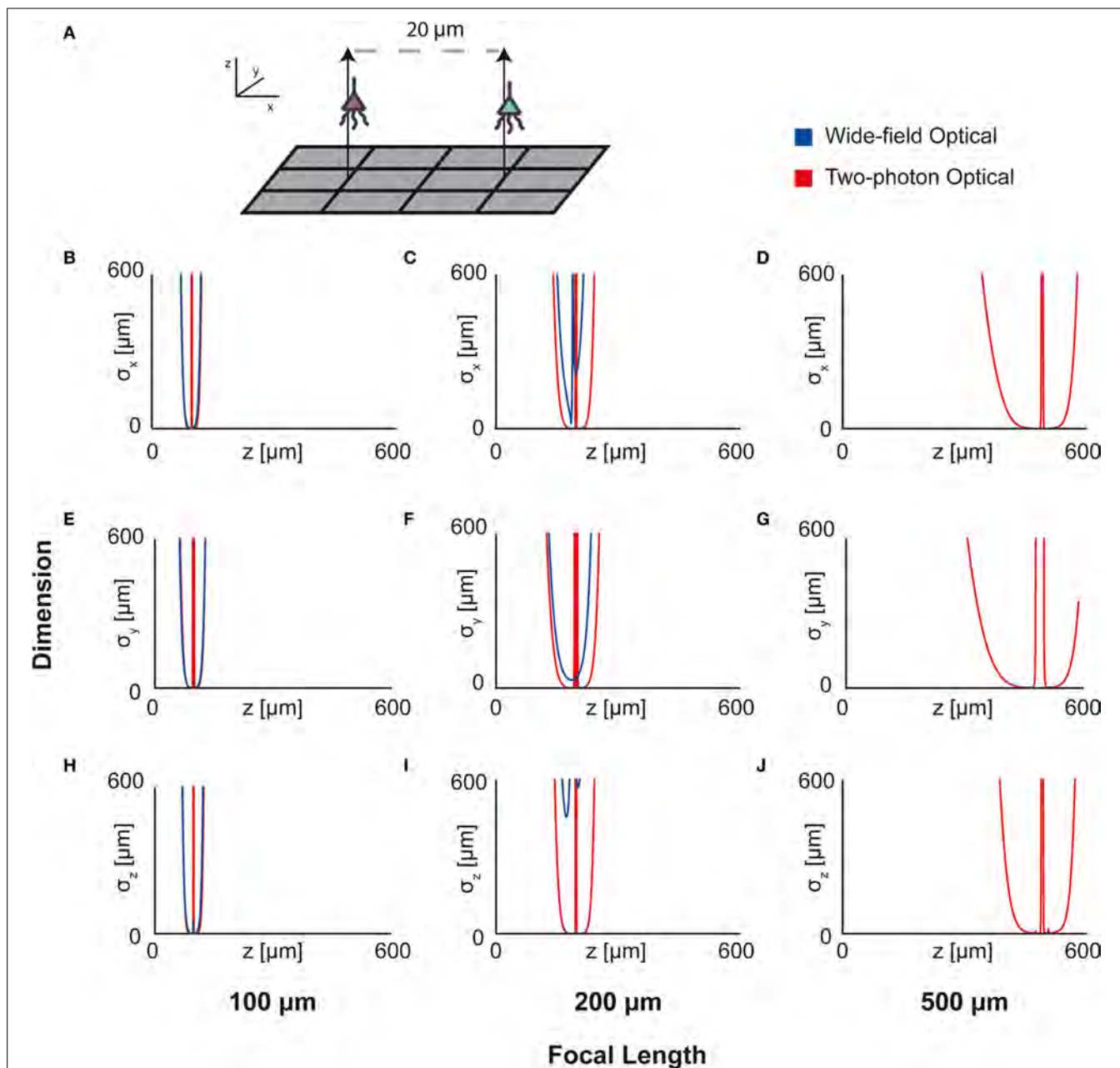


FIGURE 7 | Optical Technology Comparison at Multiple Focal Depths.

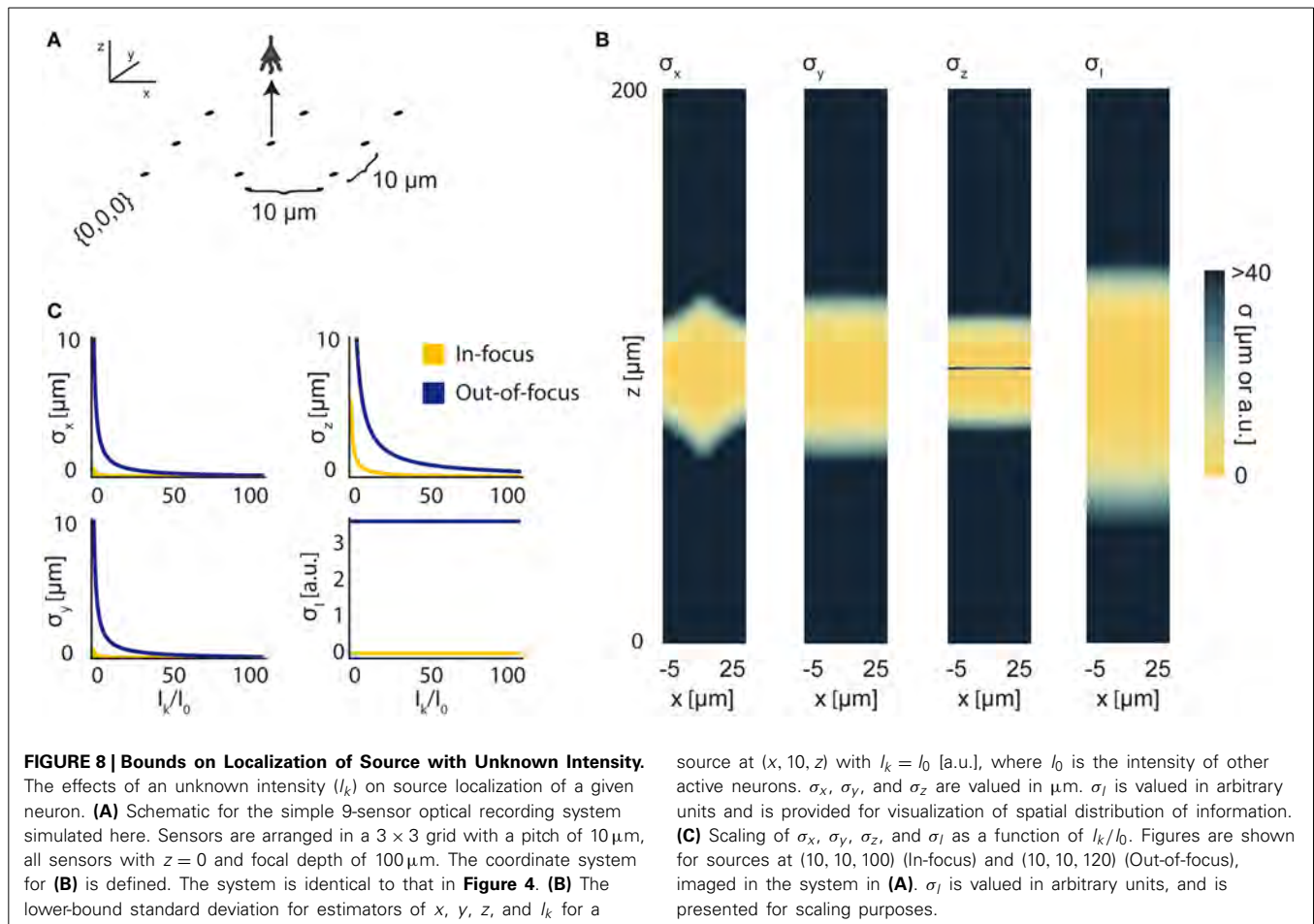
CRB on the location of the x , y , and z coordinates of a source in a multi-sensor, two-source system. The depth of the sources is varied by an equal amount and the CRB on each of the sources is calculated at each depth (the CRBs of only one source is shown; they are equivalent due to the symmetric setup). This analysis is performed for wide-field fluorescence and two-photon optical systems. (A) Schematic of recording

system: An evenly-spaced 4×3 grid of sensors detects two sources. Sensed regions have a pitch of $10 \mu\text{m}$, and neurons are separated on the x -axis by $20 \mu\text{m}$. (B,E,H) CRBs with a focal depth of $100 \mu\text{m}$. (C,F,I) CRBs with a focal depth of $200 \mu\text{m}$. (D,G,J) CRBs with a focal depth of $500 \mu\text{m}$. CRBs for the x , y , and z coordinates are in the first, second, and third rows, respectively, and are reported as standard deviations. See **Table 2** for parameter values.

using wide-field fluorescence microscopy. As in **Figure 4**, we use an array of 9 sensors in a 3×3 grid with a $10 \mu\text{m}$ pitch and attempt to localize a single source (**Figure 8A**). We simulate a $100 \mu\text{m}$ focal depth. The PSF and relevant parameters are contained in **Tables 1** and **2**, respectively. Here, we assume active background neurons have an intensity I_0 , and we are trying to

estimate the location and intensity of a neuron with unknown intensity I_k .

We find that jointly estimating intensity along with the location parameters of a source qualitatively changes the information a system carries about that source (**Figure 8B**). In comparison to a system with a fixed intensity, we find an overall decrease in



Fisher information about a source's location, as well as changes in the spatial distribution of the system's location information. As I_k increases relative to I_0 (i.e., the signal to noise ratio increases), σ_x , σ_y , and σ_z decrease (**Figure 8C**). This is largely just a restatement of our noise model: as our signal of interest outweighs background noise, it becomes easier to locate the source. The lower-bound standard-deviation of an estimator of I_k , σ_I , is invariant as I_k increases. It should be noted that our findings are contingent on our noise assumptions: should real-world noise deviate from these assumptions, the scaling properties of these results will also change.

5. DEMONSTRATIONS DISCUSSION

We have demonstrated how the Fisher information framework can be applied to neural recording technologies, and have demonstrated possible applications of this framework including determining optimal technology design and comparing technologies under differing recording conditions. In these demonstrations, interesting findings emerged, some of which confirm experimental knowledge. For instance, (1) when using columnar electrodes, increasing the spacing between electrode shanks leads to a very large fall-off in the number of neurons that can be recorded. (2) For shallow recording depths, wide-field and two-photon microscopy have similar performance capabilities, but at larger

depths two-photon microscopy becomes significantly better. (3) When the intensity of a neuron's activity is unknown, it becomes more difficult to estimate that neuron's location.

We made several simplifications regarding neural activity, noise, and recording technologies when demonstrating the use of the Fisher information framework. However, these approximations were useful in demonstrating a unifying view over recording methodologies in a single paper. Moreover, much is still experimentally unknown about noise sources and their relation to neural activity. While our demonstrations cannot give precise predictions about the capabilities of recording technologies, they demonstrate general scaling properties of the technologies, as well as illustrate situations in which the framework could be useful with more detailed models of neural recording.

A first simplification is that our demonstrations used approximate models of how neurons and noise affect sensor signals. Our demonstrations (except the last one) showed how we could use recording channels to identify the location of a fixed, known, activity. In practice these activities fluctuate over time, and can differ based on the type of neuron. As shown in our final demonstration, not knowing the intensity of neural activity worsens location estimation ability. In addition, we assumed that the effects of neural activity are linearly combined into the sensor signal. In practice, non-linear effects such as

sensor saturation may be important. Both can be incorporated into a Fisher information-based framework, although neither are treated here. Perhaps the largest simplification, the various noise sources were approximated by a simple function that ignores many potential sources of noise (see *Supplementary Material*). A comprehensive model of how noise affects neurons and sensors does not yet exist. Further research in this area will yield more informative results.

Second, we asked how we could use simplified models of recording systems to estimate the locations of neurons. For example, for optical recordings we assumed scattering through homogenous tissue, and for electrical recordings we ignored the filtering properties of electrodes. There exists a rich literature of modeling optical and electrical systems that could allow better models of recording modalities (e.g., Theer and Denk, 2006; CamuÁ-Mesa and Quiroga, 2013); incorporating these models into the framework may alleviate some of the concerns over over-simplification, and may even provide a framework for validating those models.

In order to calculate the Fisher information contained by a given technique, we need to know its PSF and noise sources. When a technology is developed, experimentally determining these functions would allow this Fisher information framework to accurately be applied. These Fisher information calculations could determine how optimal a technique's performance is. This information may then influence further design choices.

6. ADDITIONAL METHODS

6.1. NOISE CALCULATIONS

In our *Demonstrations* simulations, we make several assumptions about noise. We assume noise sources are uncorrelated (i.e., the noise from each neuron is independent and independently distributed). The sensor signal variance arises from signal dependent noise, with a standard deviation proportional to the mean signal. The signal dependent noise can be on all background neurons and/or on the sensor. As the mean activity is I_0 , the standard deviation of the activity is $\alpha \cdot I_0$, where α is a constant. The activity that reaches the sensor i (the signal) from a given neuron j then has a variance of $\sigma_j^2 = \alpha \cdot (I_0 \cdot w(\mathbf{d}^{i,j}))^2$. As the noise sources are independent, their variances can be added, so $\sigma_{noise}^2 = \sum_{j \neq k} \sigma_j^2$ (recall that we do not include noise from the neuron of interest). In simulations with two neurons of interest, we do not include noise from both neurons. We assume that neurons are uniformly distributed across the brain with density ρ_{space} and that all neurons have the same probability of firing at a given time, ρ_{fire} .

$$\begin{aligned} \sigma_{noise}^2 &= \alpha_{sensor} \rho_{fire} \rho_{space} \int_V I_0^2 w^2 dV + \alpha_{neuron} \rho_{fire} \rho_{space} \int_V I_0^2 w^2 dV \\ &= \alpha \rho_{fire} \rho_{space} \int_V I_0^2 w^2 dV \end{aligned} \quad (7)$$

In our simulations, we set $\alpha = 0.1$ (action potentials have SNRs ranging from 5 to 25, Erickson et al., 2008), $\rho_{fire} = 0.01$ (assuming neurons on average fire at 5 Hz (Harris et al., 2012) and

action potentials last ≈ 2 ms), and $\rho_{space} = 67000 \text{ mm}^3$ (dividing the number of neurons in the human brain, $\approx 8 \times 10^{10}$ (Azevedo et al., 2009) by its volume, $\approx 1200 \text{ cm}^3$ Allen et al., 2002).

6.2. DEMONSTRATIONS: ELECTRODE GRID ANALYSIS

Electrode locations were assigned to nodes on a $1 \mu\text{m}$ grid spanning a $1 \text{ mm} \times 1 \text{ mm} \times 1 \text{ mm}$ cube using the following procedures:

Columnar 6 × 6: Column locations were spaced evenly, $200 \mu\text{m}$ apart, on a 6×6 grid in the x-y plane. 101 electrodes were distributed evenly along each column, $10 \mu\text{m}$ apart.

Columnar 10 × 10: Column locations were spaced evenly, $111 \mu\text{m}$ apart, on a 10×10 grid in the x-y plane. 31 electrodes were distributed evenly along each column, $33 \mu\text{m}$ apart.

Random: Locations on the grid were drawn from a uniform random distribution with replacement.

Planar: Electrodes were placed on a uniform 61×61 grid in the x-y plane, corresponding to a grid spacing of $17 \mu\text{m}$, with a depth of $500 \mu\text{m}$.

Grid: Electrodes were placed on a uniform $15 \times 15 \times 15$ grid in the volume, corresponding to a grid spacing of $71 \mu\text{m}$.

These procedures give locations for 3636, 3100, 3636, 3721, and 3375 electrodes, respectively.

ACKNOWLEDGMENTS

We would like to thank Dario Amodei and Darcy Peterka for their helpful comments. We would like to thank Dan Dombeck for helpful discussions regarding optics and Mikhail Shapiro for discussions regarding MR applications.

Thaddeus Cybulski, Joshua Glaser, and Bradley Zamft are supported by NIH grant 5R01MH103910. Adam Marblestone is supported by the Fannie and John Hertz Foundation fellowship. Konrad Kording is funded in part by the Chicago Biomedical Consortium with support from the Searle Funds at The Chicago Community Trust. Konrad Kording is also supported by NIH grants 5R01NS063399, P01NS044393, and 1R01NS074044. George Church acknowledges support from the Office of Naval Research and the NIH Centers of Excellence in Genomic Science. Edward Boyden acknowledges funding by Allen Institute for Brain Science; AT&T; Google; IET A. F. Harvey Prize; MIT McGovern Institute and McGovern Institute Neurotechnology (MINT) Program; MIT Media Lab and Media Lab Consortia; New York Stem Cell Foundation-Robertson Investigator Award; NIH Director's Pioneer Award 1DP1NS087724, NIH Transformative Awards 1R01MH103910 and 1R01GM104948, NSF INSPIRE Award CBET 1344219, Paul Allen Distinguished Investigator in Neuroscience Award; Skolkovo Institute of Science and Technology; Synthetic Intelligence Project (and its generous donors).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fncom.2014.00172/abstract>

REFERENCES

- Aguet, F., Van De Ville, D., and Unser, M. (2005). A maximum-likelihood formalism for sub-resolution axial localization of fluorescent nanoparticles. *Opt. Express* 13, 10503–10522. doi: 10.1364/OPEX.13.010503
- Ahrens, M. B., Orger, M. B., Robson, D. N., Li, J. M., and Keller, P. J. (2013). Whole-brain functional imaging at cellular resolution using light-sheet microscopy. *Nat. Methods* 10, 413–420. doi: 10.1038/nmeth.2434
- Airy, G. B. (1835). On the diffraction of an object-glass with circular aperture. *Trans. Cambridge Philos. Soc.* 5:283.
- Alivisatos, A. P., Andrews, A. M., Boyden, E. S., Chun, M., Church, G. M., Deisseroth, K., et al. (2013). Nanotools for neuroscience and brain activity mapping. *ACS Nano* 7, 1850–1866. doi: 10.1021/nn4012847
- Allen, J. S., Damasio, H., and Grabowski, T. J. (2002). Normal neuroanatomical variation in the human brain: an mrivolumetric study. *Am. J. Phys. Anthropol.* 118, 341–358. doi: 10.1002/ajpa.10092
- Azevedo, F. A., Carvalho, L. R., Grinberg, L. T., Farfel, J. M., Ferretti, R. E., Leite, R. E., et al. (2009). Equal numbers of Neuronal and nonNeuronal cells make the human brain an isometrically scaled-up primate brain. *J. Comp. Neurol.* 513, 532–541. doi: 10.1002/cne.21974
- Bair, W., Zohary, E., and Newsome, W. T. (2001). Correlated firing in macaque visual area mt: time scales and relationship to behavior. *J. Neurosci.* 21, 1676–1697.
- Barth, A. L., and Poulet, J. F. (2012). Experimental evidence for sparse firing in the neocortex. *Trends Neurosci.* 35, 345–355. doi: 10.1016/j.tins.2012.03.008
- Broxton, M., Grosenick, L., Yang, S., Cohen, N., Andalman, A., Deisseroth, K., et al. (2013). Wave optics theory and 3-d deconvolution for the light field microscope. *Opt. Express* 21, 25418–25439. doi: 10.1364/OE.21.025418
- CamuÁ-Mesa, L. A., and Quiroga, R. Q. (2013). A detailed and fast model of extracellular recordings. *Neural Comput.* 25, 1191–1212. doi: 10.1162/NECO_a_00433
- Chaigneau, E., Wright, A. J., Poland, S. P., Girkin, J. M., and Silver, R. A. (2011). Impact of wavefront distortion and scattering on 2-photon microscopy in mammalian brain tissue. *Opt. Express* 19:22755. doi: 10.1364/OE.19.022755
- Cohen, M. R., and Kohn, A. (2011). Measuring and interpreting Neuronal correlations. *Nat. Neurosci.* 14, 811–819. doi: 10.1038/nn.2842
- Cohen, J. Y., Crowder, E. A., Heitz, R. P., Subraveti, C. R., Thompson, K. G., Woodman, G. F., et al. (2010). Cooperation and competition among frontal eye field Neurons during visual target selection. *J. Neurosci.* 30, 3227–3238. doi: 10.1523/JNEUROSCI.4600-09.2010
- Colak, S., Papaioannou, D., t Hooft, G., Van der Mark, M., Schomberg, H., Paasschens, J., et al. (1997). Tomographic image reconstruction from optical projections in light-diffusing media. *Appl. Opt.* 36, 180–213. doi: 10.1364/AO.36.000180
- Cramér, H. (1946). *Methods of Mathematical Statistics*. Princeton, NJ: Princeton University Press.
- Cremer, C., and Masters, B. R. (2013). Resolution enhancement techniques in microscopy. *Eur. Phys. J. H* 38, 281–344. doi: 10.1140/epjh/e2012-20060-1
- Dauwels, J. (2005). “Computing bayesian cramer-rao bounds,” in *Proceedings of the International Symposium on Information Theory, 2005. ISIT 2005* (Adelaide: IEEE), 425–429.
- Den Dekker, A., and Van den Bos, A. (1997). Resolution: a survey. *J. Opt. Soc. Am. A* 14, 547–557. doi: 10.1364/JOSAA.14.000547
- Denman, D. J., and Contreras, D. (2014). The structure of pairwise correlation in mouse primary visual cortex reveals functional organization in the absence of an orientation map. *Cereb. Cortex* 24, 2707–2720. doi: 10.1093/cercor/bht128
- Engelbrecht, C. J., and Stelzer, E. H. (2006). Resolution enhancement in a light-sheet-based microscope (spim). *Opt. Lett.* 31, 1477–1479. doi: 10.1364/OL.31.001477
- Erickson, J., Tooker, A., Tai, Y.-C., and Pine, J. (2008). Caged Neuron mea: a system for long-term investigation of cultured neural network connectivity. *J. Neurosci. Methods* 175, 1–16. doi: 10.1016/j.jneumeth.2008.07.023
- Gorman, J. D., and Hero, A. O. (1990). Lower bounds for parametric estimation with constraints. *IEEE Trans. Inform. Theory* 36, 1285–1301. doi: 10.1109/18.59929
- Gray, C. M., Maldonado, P. E., Wilson, M., and McNaughton, B. (1995). Tetrodes markedly improve the reliability and yield of multiple single-unit isolation from multi-unit recordings in cat striate cortex. *J. Neurosci. Methods* 63, 43–54. doi: 10.1016/0165-0270(95)00085-2
- Hampel, S., Chung, P., McKellar, C. E., Hall, D., Looger, L. L., and Simpson, J. H. (2011). Drosophila brainbow: a recombinase-based fluorescence labeling technique to subdivide neural expression patterns. *Nat. Methods* 8, 253–259. doi: 10.1038/nmeth.1566
- Harris, J. J., Jolivet, R., and Attwell, D. (2012). Synaptic energy use and supply. *Neuron* 75, 762–777. doi: 10.1016/j.neuron.2012.08.019
- Helmchen, F., and Denk, W. (2005). Deep tissue two-photon microscopy. *Nat. Methods* 2, 932–940. doi: 10.1038/nmeth818
- Helstrom, C. W. (1969). Detection and resolution of incoherent objects by a background-limited optical system. *J. Opt. Soc. Am.* 59, 164–175. doi: 10.1364/JOSA.59.000164
- Jensen, J. A. (1991). A model for the propagation and scattering of ultrasound in tissue. *J. Acoust. Soc. Am.* 89, 182. doi: 10.1121/1.400497
- Johnston, D., Wu, S. M.-S., and Gray, R. (1995). *Foundations of Cellular Neurophysiology*. Cambridge, MA: MIT press.
- Kirshner, H., Aguet, F., Sage, D., and Unser, M. (2013). 3d psf fitting for fluorescence microscopy: implementation and localization application. *J. Microsc.* 249, 13–25. doi: 10.1111/j.1365-2818.2012.03675.x
- Kording, K. P. (2011). Of toasters and molecular ticker tapes. *PLoS Comput. Biol.* 7:e1002291. doi: 10.1371/journal.pcbi.1002291
- Kullback, S. (1997). *Information Theory and Statistics*. Mineola, NY: Courier Dover Publications.
- Lambert, J. H., and Anding, E. (1892). *Lamberts Photometrie. (Photometria, sive De mensura et gradibus luminis, colorum et umbrae) (1760)*. Ostwalds Klassiker der exakten Wissenschaften. Leipzig: W. Engelmann.
- Lewicki, M. S. (1998). A review of methods for spike sorting: the detection and classification of neural action potentials. *Network* 9, R53–R78. doi: 10.1088/0954-898X/9/4/001
- Marblestone, A. H., Zamft, B. M., Maguire, Y. G., Shapiro, M. G., Cybulski, T. R., Glaser, J. I., et al. (2013). Physical principles for scalable neural recording. *Front. Comput. Neurosci.* 7:137. doi: 10.3389/fncom.2013.00137
- Marengo, E. A., Zambrano-Nunez, M., and Brady, D. (2009). “Cramer-rao study of one-dimensional scattering systems: Part i: formulation,” in *Proceedings of the 6th IASTED International Conference on Antennas, Radar, Wave Propagation (ARP’09)* (Banff, AB), 1–8.
- Matson, C., and Haji, A. (2006). *Biased Cramer-Rao Lower Bound Calculations for Inequality-Constrained Estimators (preprint)*. Air Force Research Lab, Technical report, DTIC Document.
- Mukamel, E. A., and Schnitzer, M. J. (2012). Unified resolution bounds for conventional and stochastic localization fluorescence microscopy. *Phys. Rev. Lett.* 109:168102. doi: 10.1103/PhysRevLett.109.168102
- Mukamel, E. a., Nimmerjahn, A., and Schnitzer, M. J. (2009). Automated analysis of cellular signals from large-scale calcium imaging data. *Neuron* 63, 747–760. doi: 10.1016/j.neuron.2009.08.009
- Ober, R. J., Ram, S., and Ward, E. S. (2004). Localization accuracy in single-molecule microscopy. *Biophys. J.* 86, 1185–1200. doi: 10.1016/S0006-3495(04)74193-4
- Onodera, Y., Kato, Y., and Shimizu, K. (1998). “Suppression of scattering effect using spatially dependent point spread function,” in *Advances in Optical Imaging and Photon Migration* (Orlando, FL: Optical Society of America).
- Orbach, H. S., and Cohen, L. B. (1983). Optical monitoring of activity from many areas of the *in vitro* and *in vivo* salamander olfactory bulb: a new method for studying functional organization in the vertebrate central nervous system. *J. Neurosci.* 3, 2251–2262.
- Pneumatikakis, E. A., Machado, T. A., Grosenick, L., Poole, B., Vogelstein, J. T., and Paninski, L. (2013). “Rank-penalized nonnegative spatiotemporal deconvolution-demixing of calcium imaging data,” in *Computational and Systems Neuroscience Meeting COSYNE* (Salt Lake).
- Prevedel, R., Yoon, Y.-G., Hoffmann, M., Pak, N., Wetzstein, G., Kato, S., et al. (2014). Simultaneous whole-animal 3d-imaging of Neuronal activity using light field microscopy. *Nat. Methods* 11, 727–730. doi: 10.1038/nmeth.2964
- Qin, Q. (2012). Point spread functions of the t2 decay in *k*-space trajectories with long echo train. *Magn. Reson. Imaging* 30, 1134–1142. doi: 10.1016/j.mri.2012.04.017
- Quirin, S., Pavani, S. R. P., and Piestun, R. (2012). Optimal 3d single-molecule localization for superresolution microscopy with aberrations and engineered point spread functions. *Proc. Natl. Acad. Sci. U.S.A.* 109, 675–679. doi: 10.1073/pnas.11090111108

- Sanches, J., Sousa, I., and Figueiredo, P. (2010). "Bayesian fisher information criterion for sampling optimization in asl-mri," in *2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro* (Rotterdam: IEEE), 880–883.
- Segev, R., Goodhouse, J., Puchalla, J., and Berry, M. J. (2004). Recording spikes from a large fraction of the ganglion cells in a retinal patch. *Nat. Neurosci.* 7, 1154–1161. doi: 10.1038/nn1323
- Seo, D., Carmenta, J. M., Rabaey, J. M., Alon, E., and Maharbiz, M. M. (2013). Neural Dust: An Ultrasonic, Low Power Solution for Chronic Brain-Machine Interfaces. *Q. Biol.* Available online at: arXiv:1307.2196.
- Shahram, M., and Milanfar, P. (2004). Imaging below the diffraction limit: a statistical analysis. *IEEE Trans. Image Process.* 13, 677–689. doi: 10.1109/TIP.2004.826096
- Shahram, M., and Milanfar, P. (2006). Statistical and information-theoretic analysis of resolution in imaging. *IEEE Trans. Inform. Theory* 52, 3411–3437. doi: 10.1109/TIT.2006.878180
- Shahram, M. (2005). *Statistical and Information-Theoretic Analysis of Resolution in Imaging and Array Processing*. Thesis, University of California, Santa Cruz.
- Shechtman, Y., Sahl, S. J., Backer, A. S., and Moerner, W. (2014). Optimal point spread function design for 3D imaging. *Phys. Rev. Lett.* 113:133902. doi: 10.1103/PhysRevLett.113.133902
- Shin, H.-C., Prager, R., Ng, J., Gomersall, H., Kingsbury, N., Treece, G., et al. (2009). Sensitivity to point-spread function parameters in medical ultrasound image deconvolution. *Ultrasonics* 49, 344–357. doi: 10.1016/j.ultras.2008.10.005
- Smith, R. C., and Lange, R. C. (1998). *Understanding Magnetic Resonance Imaging*. Boca Raton, FL: CRC Press.
- Theer, P., and Denk, W. (2006). On the fundamental imaging-depth limit in two-photon microscopy. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* 23, 3139–3149. doi: 10.1364/JOSAA.23.003139
- Thomann, D., Rines, D., Sorger, P., and Danuser, G. (2002). Automatic fluorescent tag detection in 3d with superresolution: application to the analysis of chromosome movement. *J. Microsc.* 208, 49–64. doi: 10.1046/j.1365-2818.2002.01066.x
- Tian, P., Devor, A., Sakadzic, S., Dale, A. M., and Boas, D. A. (2011). Monte carlo simulation of the spatial resolution and depth sensitivity of two-dimensional optical imaging of the brain. *J. Biomed. Opt.* 16:016006. doi: 10.1117/1.3533263
- Torreao, J. R., and Fernandes, J. L. (2005). "Single-image shape from defocus" in *18th Brazilian Symposium on Computer Graphics Image Processing, 2005. SIBGRAP 2005* (Natal, RN: IEEE), 241–246. doi: 10.1109/SIBGRAP.2005.47
- Van Trees, H. L. (2004). *Detection, Estimation, and Modulation Theory*. New York, NY: John Wiley & Sons.
- Winick, K. A. (1986). Cramer-rao lower bounds on the performance of charge-coupled-device optical position estimators. *J. Opt. Soc. Am. A* 3, 1809–1815. doi: 10.1364/JOSAA.3.001809
- Yan, Y., and Zeng, G. L. (2008). Scatter and blurring compensation in inhomogeneous media using a postprocessing method. *J. Biomed. Imaging* 2008, 15.
- Zador, A. M., Dubnau, J., Oyibo, H. K., Zhan, H., Cao, G., and Peikon, I. D. (2012). Sequencing the connectome. *PLoS Biol.* 10:e1001411. doi: 10.1371/journal.pbio.1001411
- Zorzos, A. N., Scholvin, J., Boyden, E. S., and Fonstad, C. G. (2012). Three-dimensional multiwaveguide probe array for light delivery to distributed brain circuits. *Opt. Lett.* 37, 4841–4843. doi: 10.1364/OL.37.004841

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 14 February 2014; accepted: 12 December 2014; published online: 21 January 2015.

Citation: Cybulski TR, Glaser JI, Marblestone AH, Zamft BM, Boyden ES, Church GM and Kording KP (2015) Spatial information in large-scale neural recordings. *Front. Comput. Neurosci.* 8:172. doi: 10.3389/fncom.2014.00172

This article was submitted to the journal *Frontiers in Computational Neuroscience*. Copyright © 2015 Cybulski, Glaser, Marblestone, Zamft, Boyden, Church and Kording. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

1 Supplementary Information

1.1 Noise Sources

The Fisher information framework allows for arbitrary noise sources, so long as they are able to be modeled. However, to demonstrate potential applications, we used a very simplified noise model that only considered signal dependent noise where the standard deviation was proportional to the mean.

There are multiple potentially relevant sources of noise that could readily be included in our model. (1) Each sensor has a constant level of noise simply due to thermal effects. (2) Many sensors have an additional variance that is proportional to the square of the signals, e.g. reference fluctuations. (3) Many sensors have an additional variance that is proportional to the signal, e.g. due to low numbers of photons (shot noise). (4,5) Each neuron may produce constant noise, e.g. background fluorescence of dyes. These neural noise sources may be independent or correlated. (6,7) Each neuron may produce variance that quadratically depends on its activation, e.g. action potentials that propagate back into varying parts of the dendritic tree.¹ These neural noise sources may be independent or correlated. (8,9) Each neuron may produce variance that linearly depends on its signal strength, e.g. fluorophore activations. These neural noise sources may be independent or correlated. We have some knowledge about the exact sizes of these signals [1], but most of these numbers are hard to know. They may be reasonable to measure in future experiments.

Taking these signals together, we obtain the following noise level on a sensor i (given a recording of N firing neurons indexed by j):

¹In a simplistic model, when a neuron fires, the action potential spreads into some variable proportion of the dendritic tree. If the recorded signal is dependent on the proportion of dendritic branches the action potential propagates into, then the standard deviation of the recorded signal is proportionate to the mean signal entering the dendrites.

$$\begin{aligned}
\sigma^2_{noise} = & \underbrace{\sigma_{sens}^2}_{\text{constant sensor noise}} + \alpha_1 \underbrace{\sum_{j=1}^N I_0^2(w(\mathbf{d}^{i,j}))^2}_{\text{ind sensor } \sigma \propto \mu \text{ SDN}} + \alpha_2 \underbrace{\sum_{j=1}^N I_0(w(\mathbf{d}^{i,j}))^2}_{\text{ind sensor } \sigma^2 \propto \mu \text{ SDN}} \\
& + \underbrace{\alpha_3 N \cdot \sigma_{neur}^2}_{\text{constant ind neuron noise}} + \underbrace{\alpha_4 N \cdot \sigma_{neur}^2}_{\text{constant corr neuron noise}} \\
& + \alpha_5 \underbrace{\sum_{j=1}^N I_0^2(w(\mathbf{d}^{i,j}))^2}_{\text{ind neuron } \sigma \propto \mu \text{ SDN}} + \alpha_6 \underbrace{\left(\sum_{j=1}^N I_0(w(\mathbf{d}^{i,j})) \right)^2}_{\text{corr neuron } \sigma \propto \mu \text{ SDN}} \\
& + \alpha_7 \underbrace{\sum_{j=1}^N I_0(w(\mathbf{d}^{i,j}))}_{\text{ind neuron } \sigma^2 \propto \mu \text{ SDN}} + \alpha_8 \underbrace{\left(\sum_{j=1}^N \sqrt{I_0(w(\mathbf{d}^{i,j}))} \right)^2}_{\text{corr neuron } \sigma^2 \propto \mu \text{ SDN}}
\end{aligned} \tag{1}$$

where *ind* and *corr* refer to independent and correlated noise sources, and *SDN* refers to signal dependent noise.

Assuming, as we do in the main text's demonstrations, that neurons are uniformly distributed and have a uniform firing rate across the entire volume:

$$\begin{aligned}
\sigma^2_{noise} = & \underbrace{\sigma_{sens}^2}_{\text{constant sensor noise}} + \alpha_1 \underbrace{\rho_{fire} \rho_{space} \int I_0^2 w^2 dV}_{\text{ind sensor } \sigma \propto \mu \text{ SDN}} + \alpha_2 \underbrace{\rho_{fire} \rho_{space} \int I_0 w dV}_{\text{ind sensor } \sigma^2 \propto \mu \text{ SDN}} \\
& + \underbrace{\alpha_3 \rho_{fire} \rho_{space} V \cdot \sigma_{neur}^2}_{\text{constant ind neuron noise}} + \underbrace{\alpha_4 \rho_{fire} \rho_{space} V \cdot \sigma_{neur}^2}_{\text{constant corr neuron noise}} \\
& + \alpha_5 \underbrace{\rho_{fire} \rho_{space} \int I_0^2 w^2 dV}_{\text{ind neuron } \sigma \propto \mu \text{ SDN}} + \alpha_6 \underbrace{\rho_{fire} \rho_{space} \left(\int I_0 w dV \right)^2}_{\text{corr neuron } \sigma \propto \mu \text{ SDN}} \\
& + \alpha_7 \underbrace{\rho_{fire} \rho_{space} \int I_0 w dV}_{\text{ind neuron } \sigma^2 \propto \mu \text{ SDN}} + \alpha_8 \underbrace{\rho_{fire} \rho_{space} \left(\int \sqrt{I_0 w} dV \right)^2}_{\text{corr neuron } \sigma^2 \propto \mu \text{ SDN}}
\end{aligned} \tag{2}$$

Both constant and shot noise terms can be minimized in their effect by optimizing the experimental design, e.g. through good dyes and strong illumination (but see [1]).

In addition, in the main text we assume that the noise is Gaussian, which has also been assumed in previous statistical formulations [2, 3]. This assumption has been shown to be valid for thermal noise and shot noise in some conditions [4, 5].

1.2 Fisher Information Shortcut Derivation

We have assumed that the distribution of signals on a sensor is a Gaussian distribution: $f(X; \boldsymbol{\theta}) = \mathcal{N}(\mu(\boldsymbol{\theta}), \sigma_{noise}^2)$, where μ is a function that describes the mean signal as a function of the parameters in $\boldsymbol{\theta}$. In practice, $\boldsymbol{\theta}$ represents the positions and intensities of the various neurons being analyzed, and $\mu(\cdot)$ is some representation of the PSFs that relate signals to sensors. Due to this Gaussian assumption, solving for the Fisher information matrix simplifies:

$$\mathcal{I}_{jk} = E \left[\left(\frac{\partial}{\partial \theta_j} \ln(f(X; \boldsymbol{\theta})) \right) \left(\frac{\partial}{\partial \theta_k} \ln(f(X; \boldsymbol{\theta})) \right) \right] \quad (3)$$

$$\begin{aligned} \frac{\partial}{\partial \theta_j} \ln(f(X; \boldsymbol{\theta})) &= \frac{1}{f(X; \boldsymbol{\theta})} \frac{\partial}{\partial \theta_j} f(X; \boldsymbol{\theta}) \\ &= \frac{1}{\frac{1}{\sqrt{2\pi\sigma_{noise}^2}} \exp\left(\frac{-(X-\mu(\boldsymbol{\theta}))^2}{2\sigma_{noise}^2}\right)} \\ &= \frac{1}{\sqrt{2\pi\sigma_{noise}^2}} \exp\left(\frac{-(X-\mu(\boldsymbol{\theta}))^2}{2\sigma_{noise}^2}\right) \cdot \\ &\quad \frac{\partial}{\partial \theta_j} \left(\frac{-(X-\mu(\boldsymbol{\theta}))^2}{2\sigma_{noise}^2} \right) \\ &= \frac{1}{2\sigma_{noise}^2} \cdot 2(X-\mu(\boldsymbol{\theta})) \cdot \frac{\partial \mu(\boldsymbol{\theta})}{\partial \theta_j} \\ &= \frac{(X-\mu(\boldsymbol{\theta}))}{\sigma_{noise}^2} \frac{\partial \mu(\boldsymbol{\theta})}{\partial \theta_j} \end{aligned} \quad (4)$$

$$\begin{aligned}
\mathcal{I}_{jk} &= E \left[\frac{(X - \mu(\boldsymbol{\theta}))^2}{\sigma_{noise}^4} \frac{\partial \mu(\boldsymbol{\theta})}{\partial \theta_j} \frac{\partial \mu(\boldsymbol{\theta})}{\partial \theta_k} \right] \\
&= E \left[\frac{(X - \mu(\boldsymbol{\theta}))^2}{\sigma_{noise}^4} \right] \frac{\partial \mu(\boldsymbol{\theta})}{\partial \theta_j} \frac{\partial \mu(\boldsymbol{\theta})}{\partial \theta_k} \\
&= \frac{1}{\sigma_{noise}^2} \frac{\partial \mu(\boldsymbol{\theta})}{\partial \theta_j} \frac{\partial \mu(\boldsymbol{\theta})}{\partial \theta_k}
\end{aligned} \tag{5}$$

2 References

- [1] Adam H Marblestone, Bradley M Zamft, Yael G Maguire, Mikhail G Shapiro, Thaddeus R Cybulski, Joshua I Glaser, Dario Amodei, Ben Stranges, Reza Kalhor, David A Dalrymple, Dongjin Seo, Elad Alon, Michel M Maharbiz, Jose M Carmena, Jan M Rabaey, Edward S Boyden, George M Church, and Konrad P Kording. Physical principles for scalable neural recording. *Front Comput Neurosci*, 7:137, 2013.
- [2] Morteza Shahram. *Statistical and information-theoretic analysis of resolution in imaging and array processing*. Thesis, 2005.
- [3] Morteza Shahram and Peyman Milanfar. Statistical and information-theoretic analysis of resolution in imaging. *Information Theory, IEEE Transactions on*, 52(8):3411–3437, 2006.
- [4] Robert Tyson. *Adaptive optics engineering handbook*, volume 67. CRC Press, 1999.
- [5] Alan C Bovik. *Handbook of image and video processing*. Access Online via Elsevier, 2010.