

An information-theoretic view of EEG sensing

Pulkit Grover and Praveen Venkatesh, Carnegie Mellon University
(pulkit@cmu.edu, praveen1@andrew.cmu.edu)

Abstract

This paper uses an information-theoretic lens to examine the use and implementation of Electroencephalography (EEG) systems for neural imaging. There is a widespread belief in the clinical and neuroscientific community that ultra high-density EEG imaging of the brain (beyond the typically used few hundred or fewer electrodes) will not yield higher spatial resolution. Theoretically, this belief rests on a spatial Nyquist rate analysis on idealized human head models, which relies on the understanding that high-spatial frequencies (that carry high resolution information) decay as they travel from the brain to the scalp surface. Interestingly, the belief continues to be held despite being recently challenged experimentally, in part because of the spatial Nyquist results. In this work, we question the spatial Nyquist rate analysis, mathematically as well as conceptually. Mathematically, we correct and generalize the derivation of the spatial transfer function for the brain-to-scalp signal propagation as well as the ensuing Nyquist-rate analysis, and then use it to obtain improved estimates for spatial Nyquist rates. Conceptually, we observe that the Nyquist-rate analysis provides a limit on the following problem: what is the minimum number of sensors required to reconstruct the scalp EEG with a specified (small) mean-squared error? However, this problem is fundamentally different from the imaging problem of minimizing error in reconstructing the *neural source* inside the brain, which requires an understanding of noise along with signal. To that end, we utilize the transfer function to obtain an information-theoretic fundamental limit on the achievable accuracy in localizing single-dipole sources. The technique relies on computing the distortion-rate function for a single dipole source and evaluating it at an upper bound on mutual information across the brain-to-scalp channel, and can be extended to more general sources as well. Finally, we observe that the main obstacle in understanding the required number of EEG sensors is the lack of ultra-high-density systems that can test these limits, and information theory can prove helpful in this context as well. One engineering difficulty is that the required circuit area and energy for sampling EEG signals can be too large to enable these systems be safe, compact and portable. Towards reducing this area and energy, we observe that the decay of high spatial frequencies in EEG, while being a detriment to reconstruction accuracy (which it is), also leads to large the spatial correlations in recorded EEG signals. Interestingly, these correlations can be harnessed using a “hierarchical referencing” technique to provide a novel information-theoretic hierarchical referencing technique that reduces circuit energy and area to enable high-density implementations. We survey some of our recent numerical and experimental results, and further show that this technique requires order-sense fewer number of bits than the classical information-theoretic strategies.

I. INTRODUCTION AND MOTIVATION

Electroencephalography (EEG) systems sense brain activity by measuring electric potentials generated by the brain on the scalp surface, typically using a few hundred or fewer electrodes. EEG is not commonly used as an imaging modality, and is instead used to make simpler inferences. A competing noninvasive modality, Magnetoencephalography (MEG), relies on measuring magnetic fields, and also uses only a few hundred sensors. In comparison, MEG “SQUID” magnetometers (superconducting quantum interference devices) are much more expensive¹, and restrict the subjects’ movements. Despite these shortcomings, why is MEG used as an imaging modality, but not EEG? This is because MEG is thought to provide higher resolution for brain-activity localization and imaging², although it is unclear if EEG is fundamentally limited to a resolution below that of MEG, or whether an increase in number of electrodes possibly with a reduction in electrode noise, can make EEG’s resolution comparable with, or even exceed, that of MEG. The spatial resolution of both these technologies is of course lower than that of invasive technologies such as Electrocorticography (ECoG; where electrodes are placed on the surface of the brain after a surgery) or multielectrode arrays (MEAs), but also substantially lower than another noninvasive modality: functional Magnetic Resonance Imaging (fMRI). However, fMRI measures the hemodynamic response to transient neural activity, which evolves slowly over time, thus reducing the temporal resolution to a few seconds. On the other hand, EEG and MEG easily offer temporal resolution of hundreds of microseconds. This raises interesting tradeoffs in using different (invasive and noninvasive) modalities, in conjugation or in isolation. The question of limits of these modalities has been considered in the past (e.g. [5]), and is garnering increased attention from the community [6] due to advances in circuit and device designs. The goal of this paper is to cast an information-theoretic lens on the problem of noninvasive neural recording and imaging using EEG, with the hope that similar approaches can be adopted to examine other modalities (especially those that rely on electromagnetic signals for imaging, whether invasive or noninvasive). We will explore fundamental approaches to understand limits of inferring information from such recordings, and also discuss how information-theoretic techniques can prove useful in reducing energy and volume requirements on these systems.

One reason why EEG is thought to only provide low spatial resolution is that the high spatial frequency content of EEG signals is severely attenuated (relative to low spatial frequency content; see Fig. 1). It is these high frequencies that carry high spatial resolution information. This spatial low-pass filtering effect of the human head has been used to estimate the “spatial Nyquist rates” of EEG using idealized head models by Srinivasan et al. [2], [7]³. It appears to us that this Nyquist-rate analysis, which concludes that 128

¹For example, the entire western Pennsylvania has only one MEG system at the University of Pittsburgh Medical Center, and is shared for research and clinical purposes.

²There are studies that observe that the localization accuracy (e.g. [1] and references therein) for the the two modalities are comparable, or complementary [2, Ch. 2], and sometimes, EEG is noted to be better [3]. However, in practice, EEG is not often used for imaging, but instead for characteristic signatures such as Event-Related Potentials (ERPs) or emissions in temporal frequency bands that are well established biomarkers [4].

³See [2] for an excellent introduction to the models and techniques.

sensors achieve the Nyquist rate⁴, along with practitioners' observations of low-spatial resolution of existing EEG systems (which also use a few hundred electrodes), has lead to a widespread belief that the resolution provided by EEG saturates at a few hundred electrodes.

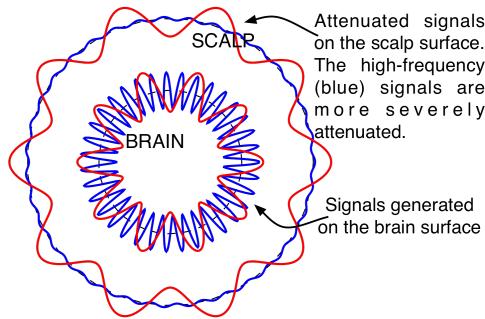


Fig. 1. A qualitative illustration of the low-pass filtering effect of the brain-to-scalp channel on an idealized spherical head with toy low and high frequency signals illustrated on a spherical approximation of the head. A low-spatial-frequency signal (red) has a substantially smaller decay than a high-spatial frequency signal (blue), even when their relative amplitudes are comparable.

The first part of this paper examines this belief through theoretical and simulation-based analyses, as well as examining existing experimental evidence that indicates that the belief is incorrect. This research has motivated an ongoing project for building the “Neural Web”: an ultra-high-density portable EEG system with thousands of electrodes to test the ultimate limits of EEG, and this paper brings out how information-theoretic concepts and analyses are crucial enablers in conceptualizing and implementing this system. Concretely, the contributions of this paper are:

- 1) The first EEG sensing problem addressed is recovery of the scalp EEG signal to within a small mean-squared error over the entire scalp. We correct the spatial Nyquist rate analysis presented in [2], [7] for required spatial sampling for EEG signal reconstruction. The correction is at two levels (i) derivation of the transfer function of the signal at the surface of the human brain as it passes through various layers and is measured on the scalp; and (ii) in the ensuing computation of the spatial Nyquist rate using the transfer function.
- 2) The second EEG sensing problem addressed is the recovery of the *source of activity* of the EEG signal. Using the transfer function obtained to provide an expression for fundamental limits on EEG's source-localization accuracy for a single dipole source inside idealized head models and for EEG signals measured at a given time instant (assuming a uniform distribution of the activity in the brain). These limits are evaluated for simplified noise models. Our work goes beyond the existing lower bounds of Mosher *et al.* [5] on source-localization accuracy for EEG. Their bounds only hold for unbiased estimators because they are based on Cramer-Rao bounds [8]. This is a limitation because biased estimators can outperform optimal unbiased estimators (and thus Cramer-Rao bounds) in mean-squared error sense (e.g. [9], [10]), and even for the particular case of EEG source localization, commonly used algorithms such as the Minimum Norm Estimate [11] have a well-known bias towards the brain surface [12]. Our information-theoretic bounds hold for all source-localization algorithms, unbiased as well as biased. Further, the approach we present can be extended to understand effects of change in sensor density and number, placement of electrodes, noise in electrodes, etc., using tools from information theory;
- 3) The third is an engineering problem in acquiring EEG signals. We survey and advance on our recent work on exploiting the low-pass nature of this transfer function to reduce energy consumption in circuitry of EEG systems. While the decay of high spatial frequencies is conventionally viewed as a detriment to EEG resolution (which it is), we harness spatial correlations induced by this decay to provide a novel information-theoretic hierarchical referencing technique that reduces circuit energy and area to enable high-density implementations.

In Section III-A, we first correct and improve the analysis for idealized head models in [2], [7], concluding that these works underestimate the spatial Nyquist rate of EEG signals, but that our revised Nyquist-rate estimate is comparable to the number of electrodes used by modern systems. Interestingly, recent experimental evidence strongly indicates that there is information to be gained by sampling signals at inter-sensor distances of a few mm [13]–[16], which will require thousands of electrodes for whole scalp coverage (see Fig. 9). Thus, these results strongly suggest that sampling dramatically beyond the spatial Nyquist rate could still be useful. An immensely interesting question is: why do the experimentalists observe an advantage when they go beyond the spatial Nyquist rate? One might think that this is because of the idealistic assumptions (e.g. zero noise and spherical head models) made in the Nyquist rate analysis. However, in Section III-B, we question⁵ whether the EEG spatial Nyquist-rate formulation is at all relevant to fundamental limits on required spatial sampling. In particular, when the goal is not simply to recover the scalp EEG signal, but to process the scalp EEG to reconstruct the signal source *inside the brain*, Nyquist-rate bounds can be very loose. This is because the Nyquist-rate formulation attempts to minimize the mean-squared error in reconstructing the EEG signal *on the scalp surface*, not the error in reconstructing the signal source *inside the brain*. We believe that this is the reason that recent evidence [13]–[16] report

⁴We note that a careful reading of Srinivasan *et al.* [7] makes it clear that their goal was to *increase* the number of electrodes used in EEG by providing a *lower*, not upper, bound.

⁵We note that the beginnings of this line of questioning can be seen in the conclusions section of [7].

substantial advantages to going significantly beyond the spatial Nyquist rate: there is still high-frequency spatial information in EEG, but it does not help improve mean-square errors in EEG reconstruction because it is represented in small amplitude signals. Our simulation results (Section III-B) also show that the source-localization accuracy also increases substantially from a few hundred to a few thousand electrodes, suggesting that even a few thousand electrodes may not be sufficient for close to optimal source localization. Can information theory provide a framework for understanding fundamental limits on EEG source-localization? In Section III-C, recognizing that the brain-to-scalp transfer function characterizes a “communication channel” through which a source can be viewed as attempting to communicate its location, we use information-theoretic tools to derive fundamental limits on the accuracy of EEG signals. The derivation of the spatial transfer function is helpful here: we utilize the spatial transfer function to obtain a bound on the Shannon capacity of this channel when a single dipole “uses the channel” to “communicate its location” to electrodes on the scalp. This is turned around to obtain a lower bound on the mean-squared error in reconstructing the dipole’s location through classical rate-distortion techniques.

Finally, information theory is also helpful in guiding system design. The system design problem is nontrivial because state-of-the-art EEG recording systems require 5 mW of power per electrode (e.g. [17]), which quickly scales to tens of watts for thousands of electrodes for *just the sensing circuitry*. There will be additional power costs for control logic and communication circuitry. Such power levels will be too large to allow the system to be portable, or even to keep the scalp perspiration-free, and in the worst case, the system may be harmful for moderate-term use. In Section IV we observe that the decay of high spatial frequencies in EEG, while a detriment to signal quality overall, also provides an opportunity for reducing circuit area and power in high-density sensing. This is because decay of these frequencies induces large spatial correlations in nearby electrodes (see Fig. 13) that can be harnessed through information-theoretic techniques. In particular, we provide a novel hierarchical referencing mechanism that reduces power consumption of Analog-to-Digital Converters (ADCs; a dominant power sink) by 3x or more (in idealized head models as well as extrapolated experimental data; with conservative assumptions on power savings) in comparison with classical strategies. While the problem can be viewed as one of compressing a Gauss-Markov source [18]–[20], which has been widely studied, this novel strategy is needed because classical lossy compression strategies for Markov sources, e.g., Differential Pulse-Coded Modulation (DPCM) [20], are not applicable here⁶. Building on some of our recent work in this direction, here we show that the provided strategies can outperform classical approaches in order sense (in total number of ADC bits). We conclude in Section V.

While the results of this paper do not conclude that ultra-high-density EEG systems will definitely yield substantially higher resolution than currently believed, they do point out that the believed limits to EEG resolution are pessimistic, and that an informational-line of reasoning offers an explanation for the discrepancy between recent experimental results and currently accepted fundamental limits. Indeed, one needs to bring together theory and experiments in order to test when the spatial resolution saturates. Without such attempts, bulky, non-portable, and expensive MEG systems will remain the state-of-the-art in dynamic noninvasive brain recording and Brain Computer Interfaces (BCIs), and the full capabilities of EEG systems will never be realized. Further, we observe that information theory in particular, and science of information more broadly, have a crucial role to play in not only obtaining the fundamental limits to these systems, but also in guiding their design and implementation.

II. SPHERICAL HARMONICS, SPHERICAL HEAD MODELS, NOTATION AND PRELIMINARIES

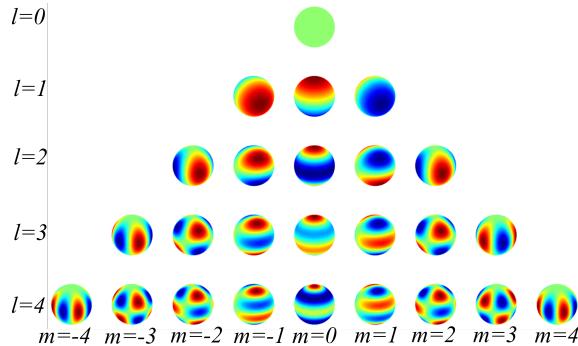


Fig. 2. The real part of the spherical harmonic basis functions $Y_{l,m}$ ’s for $l = 0, 1, 2, 3, 4$.

We refer the reader to [21, Chapter 19] for an introduction to spherical harmonics. A suitably nice function $f(\theta, \phi)$ on the surface of a sphere, where θ is the polar angle, and ϕ is the azimuthal angle (in spherical coordinates), can be represented as a linear combination of *spherical harmonics* $Y_{l,m}$ ’s of degree $l \geq 0$ and order $m = -l, -l + 1, \dots, l - 1, l$:

$$f(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l c_{l,m} Y_{l,m}(\theta, \phi), \quad (1)$$

⁶Our distributed sensing problem imposes the constraint that the reconstruction of one sensor’s measurement at the receiver is not available at another, which is needed in DPCM.

where $Y_{l,m}(\theta, \phi) = \sqrt{\frac{(2l+1)}{4\pi} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos \theta) e^{im\phi}$, are orthonormal basis functions ($\int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} |Y_{l,m}(\theta, \phi)|^2 \sin \theta d\theta d\phi = 1$), and $P_l^m(\cdot)$ is the Legendre polynomial of the first kind [21]. Larger values of l correspond to higher spatial frequencies. E.g., $l = 0$ is the lowest spatial frequency corresponding to the “DC-value” (i.e., spatial average over the entire sphere) of the signal.

Lemma 1 (Parseval’s relation for spherical harmonics): For normalized spherical harmonics $Y_{l,m}$, Parseval’s relation is given by $\int_{\mathbf{x} \in \partial\mathcal{S}} |f(\mathbf{x})|^2 d\mathbf{x} = \sum_{l=0}^{\infty} \sum_{m=-l}^l |c_{l,m}|^2$, where $c_{l,m}$ are the spherical harmonics coefficients of $f(\cdot)$, and $\partial\mathcal{S} = \{\mathbf{x} : \|\mathbf{x}\|_2 = r\}$ denotes the boundary of sphere \mathcal{S} of radius r .

Because of the orthonormality of the spherical harmonics basis function, we can define the “energy” contained in harmonics $l = l_1, l_1 + 1, \dots, l_2$ as $E_{l_1}^{l_2} = \sum_{l=l_1}^{l_2} \sum_{m=-l}^l |c_{l,m}|^2$.

A. The “ N -sphere model” of the head for computing potentials

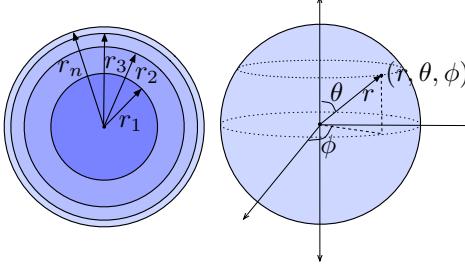


Fig. 3. The N -sphere model. Conductivities σ_i of different layers are different. Also shown are the polar and azimuthal angles (θ and ϕ).

The N -sphere model of the human head is a simplistic model that frequently serves as a first step in understanding EEG signals because it is easier to analyze mathematically than more realistic models [2]. The innermost sphere is the brain, and the outermost sphere is the scalp. A common approximation is the 4-sphere model with brain, cerebrospinal fluid (CSF), skull, and scalp layers. The layers have radii $r_1 = 7.8$ cm, $r_2 = 7.9$ cm, $r_3 = 8.6$ cm, and $r_4 = 9.2$ cm [2]. The conductivities of the layers have had different estimates over the years. Typical reported numbers of conductivity for the various layers (denoted by σ_i), with respect to the conductivity of the brain (denoted by σ_1) are as follows: $5\sigma_1$ for CSF, $\sigma_1/15$ for the skull, and σ_1 (equal to brain conductivity) for the scalp [2], [22]. The conductivity of the skull (σ_3) is particularly controversial. While Nunez and Srinivasan [2] use $\sigma_3/\sigma_1 = 1/80$, recent estimates find that skull conductivity is higher, and ratios of $1/15$ are now well accepted [22], with the understanding that there is variation between individuals. Therefore, we use a range of values from $1/80$ to $1/15$ to illustrate the range of Nyquist rates.

Solving Laplace’s equation in spherical coordinates for this model, the potential in the spherical annuli between $(s-1)$ -th and s -th sphere (denoted generally by $\Phi^{(s)}$) is given by [2]:

$$\Phi^{(s)}(r, \theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l \left(A_{lm}^{(s)} \left(\frac{r}{r_s} \right)^l + B_{lm}^{(s)} \left(\frac{r_s}{r} \right)^{l+1} \right) Y_{lm}(\theta, \phi), \quad \text{for } r_{s-1} \leq r \leq r_s \quad (2)$$

so that at $r = r_s$, the field is: $\Phi^{(s)}(r_s, \theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l \left(A_{lm}^{(s)} + B_{lm}^{(s)} \right) Y_{lm}(\theta, \phi)$. As a reminder: here, $\Phi^{(s)}$ is the expression for the potential *inside* the s -th shell.

The following section builds on well-known solutions for Laplace’s equation with spherical boundary conditions (see, e.g., [2], [21]).

III. INFORMATION-THEORETIC LENS ON SPATIAL SAMPLING FOR EEG IMAGING

We first compute the brain-surface-to-scalp spatial transfer function in the spherical harmonic basis for the idealized N -sphere model. We then use the transfer function to compute spatial Nyquist rates and obtain fundamental limits on source localization accuracy for human EEG.

A. The spatial transfer function and Nyquist rates for EEG

1) *Solving the N -sphere model to compute the spatial transfer function:* We refer the reader to [21] [2, Appendix I] for a quick review of spherical harmonics. The following theorem provides the spatial transfer function that can be used to compute the potential at the surface of the outermost sphere for a given potential at the surface of the innermost sphere (specified by coefficients $c_{l,m}^{(1)}$ of spherical harmonics $Y_{l,m}$ ’s; see (1)), given the conductivities and the radii of the intermediate shells. For EEG, we can use this to compute the potential on an idealized scalp surface given the potential on the idealized brain surface.

Theorem 1 (Spatial transfer function): For the N -sphere model described in Section II-A, the following recursive equations describe the spatial transfer function connecting $c_{l,m}^{(N)}$ and $c_{l,m}^{(1)}$, where $c_{l,m}^{(N)} = A_{lm}^{(N)} + B_{lm}^{(N)}$, and $c_{l,m}^{(1)} = A_{lm}^{(1)} + B_{lm}^{(1)}$. Here, $B_{lm}^{(s)} = \gamma_l^{(s)} A_{lm}^{(s)}$, for $s = 1, 2, \dots, N$ and A_{lm} ’s described through a recursive relationship:

$$A_{lm}^{(s)} = \frac{A_{lm}^{(s-1)} + B_{lm}^{(s-1)}}{\left(\left(\frac{r_{s-1}}{r_s} \right)^l + \gamma_l^{(s)} \left(\frac{r_s}{r_{s-1}} \right)^{l+1} \right)}, \quad (3)$$

for $s = 2, 3, \dots, N$. Here, $\gamma_l^{(s)} = \frac{\frac{l}{l+1}\sigma_{s,s+1} - \zeta_{l,s+1}}{\sigma_{s,s+1} + \zeta_{l,s+1}}$, for $s = 2, 3, \dots, N-1$ and $\gamma_l^{(N)} = \frac{l}{l+1}$, where $\zeta_{l,s} = \frac{l\left(\frac{r_{s-1}}{r_s}\right)^l - (l+1)\gamma_l^{(s)}\left(\frac{r_s}{r_{s-1}}\right)^{l+1}}{\left(\frac{r_{s-1}}{r_s}\right)^l + \gamma_l^{(s)}\left(\frac{r_s}{r_{s-1}}\right)^{l+1}}$, and $\sigma_{s,s+1} = \frac{\sigma_s}{\sigma_{s+1}}$.

Proof: This result extends and corrects the analysis in [2], [7]. See Appendix A for a detailed proof. \blacksquare

Remark: Because the expression for this transfer function can be useful in many analyses and even in forward model generation in the N -sphere model, we now specify which expressions and analyses in [2] are erroneous (in the notation adopted in [2]): (i) The expression (G.2.1) on pg. 563 in [2] should have $\frac{n}{n+1}$ instead of $\frac{n+1}{n}$ in the first term in the numerator (we note that the authors use n instead of l here as an index for spherical harmonic degree). This can result in substantial error in the transfer function at low frequencies; (ii) The expression in (G.2.2) should be $\frac{r_{23}^n \frac{n}{n+1} - V_n r_{32}^{n+1}}{r_{23}^n + V_n r_{32}^{n+1}}$ instead of $r_{23} \frac{\frac{n}{n+1} - V_n r_{32}^{n+1}}{r_{23}^n + V_n r_{32}^{n+1}}$. The second error appears to us to be a typographical error, while the first error appears to be an error in algebraic manipulations. (iii) Finally, a third error is in the spatial Nyquist-rate estimation (we discuss this in further detail in Section III-A): the authors do not add up the energy in the $(2l+1)$ harmonics (for a fixed l) in their estimation of energy in the l -th harmonic and instead just use the transfer function directly to estimate Nyquist rates. This error is most easily⁷ observed through comparing the two plots in Fig. 4 in [7]: the Nyquist rate required is estimated by obtaining (through visual inspection) the number of sensors below which aliasing due to a finite number of electrodes starts affecting the transfer function. As shown in Figs. 5 and 6, the total energy (summed over all m for each l) is very different from the transfer function. These errors are nevertheless important and difficult to catch because no proofs are provided in the manuscript or papers. This theorem thus corrects and generalizes these expressions. For readers intending to utilize our expressions or derivation, we provide a detailed proof in Appendix A, and a freely downloadable MATLAB source code at [24].

The precision of the approximation is illustrated in Fig. 4.

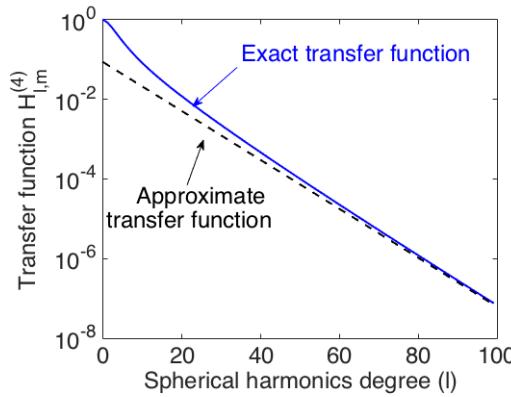


Fig. 4. A comparison of the exact and the approximate transfer functions (obtained in Theorem 1 and Corollary 1). As expected, the approximation is increasingly precise at high frequencies (*i.e.*, large l). The parameter values for this plot are $r_1 = 7.9$ cm, $r_2 = 8.0$ cm, $r_3 = 8.6$ cm, and $r_4 = 9.1$ cm, with conductivities for the 4 layers being 1, 5, 1/15, and 1 respectively.

In [2], the authors provide numerical approximations to fields generated by radial dipoles that are convenient to use in practice [2, Pg. 565]. Here, we instead simplify the expressions in Theorem 1 to provide different approximations that hold in the limit of large l . These approximations provide intuition into the low-pass filtering nature of the brain-to-scalp transfer function:

Corollary 1 (Exponential decay with spatial frequency): For large values of l , the transfer function can be approximated as:

$$H^{(N)}(l, m) = \frac{c^{(N)}(l, m)}{c^{(1)}(l, m)} \approx \left(\prod_{s=2}^N \frac{2\sigma_s}{\sigma_{s+1} + \sigma_s} \right) \frac{1}{\left(\frac{r_N}{r_1}\right)^{l+1}} \quad \text{with } \sigma_{N+1} = 0. \quad (4)$$

Proof: See Appendix B. \blacksquare

Remark: Notice that higher spatial frequency terms decay primarily because of the term $\left(\frac{r_N}{r_1}\right)^{l+1}$ in the denominator, which is the reason why the decay is exponential in l (as noted earlier, larger values of l correspond to higher frequencies). This approximation shows that the transfer function is described quite accurately by a term that depends simply on the distance from the brain surface (up to a scaling constant). In the existing literature, skull's poor conductivity (and its mismatch with conductivities of neighboring layers) is often held responsible for the low-pass effect. Indeed, skull's conductivity does have a low-pass filtering effect that is missed by this approximation. This additional low-pass filtering effect can be seen in Fig. 6, where the impact of different conductivities on Nyquist rate is studied. If skull's conductivity did not have an additional low-pass filtering effect, all the curves in Fig. 6 would look identical. Nevertheless, for large l , the effect of distance dominates, as can be inferred from Fig. 4. We also note that the transfer function does not depend on σ_1 . This might seem surprising. The reason σ_1 does not appear in the expression is because

⁷In another work [23], the authors explicitly say “*The spatial spectra of scalp potential corresponding to a white noise cortical source distribution has already been calculated as shown in figure 2,*” whereas Fig. 2 in [23] only contains the spatial transfer function.

by specifying $c_{l,m}^{(1)}$ for all l, m , we specify the potential ϕ on the sphere of radius r_1 . With this boundary condition, the conductivity σ_1 does not matter. However, to compute the potential due to a source inside the brain (e.g. a dipole source used in Lemma 2 in Section III-C) requires computing the $c_{l,m}^{(1)}$ values, which require use of σ_1 .

2) *Nyquist rates for a 4-sphere model: Analysis and numerical results:* Because the transfer function never becomes zero, one can think of measuring the spatial bandwidth at a threshold that contains (say) 99% of the signal's energy. This philosophy has been implicitly adopted in the work of Srinivasan et al. [2], [7] where the authors visually examine the transfer function [7, Fig. 4] and conclude that 128 sensors are sufficient for reconstructing the EEG.

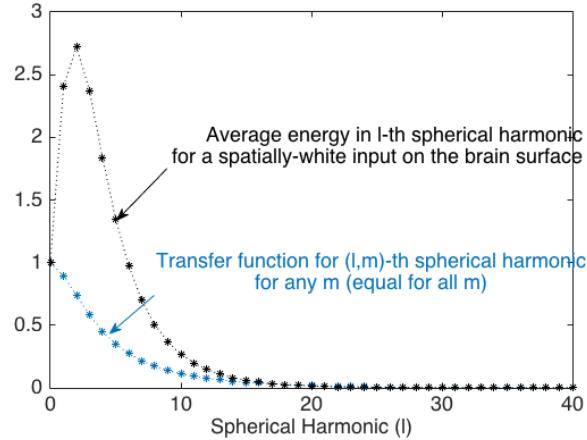


Fig. 5. Spatial transfer function as a function of l (equal for all m , see Appendix A). The transfer function decays exponentially with l . Also shown is the average energy contained in l -th spatial frequency ($\sum_{m=-l}^l |H_{l,m}^{(4)}|^2 = (2l+1)|H_{l,m}^{(4)}|^2$ because $H_{l,m}^{(4)}$ does not depend on m) as l varies for a spatially white source on the cortical surface (i.e., average $|c_{l,m}^{(1)}|^2 = 1$ at the brain surface). Energy initially increases as a function of l because of the $(2l+1)$ factor, but eventually the exponential decay takes over. This factor is ignored in [7], which is one of the reasons for their low estimates of spatial Nyquist rates. Code available at [24].

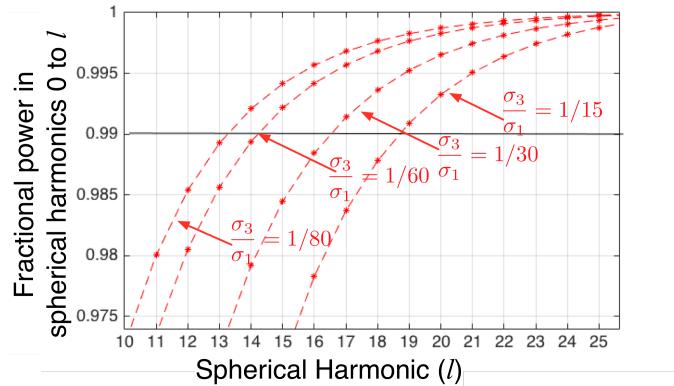


Fig. 6. To capture 99% of the energy requires sampling up to $l = 19$ in the worst case ($\frac{\sigma_3}{\sigma_1} = 1/15$), which corresponds to 361 degrees of freedom. Even for $\frac{\sigma_3}{\sigma_1} = 1/30$, the Nyquist rate is $l_0 = 17$, which has $17^2 = 289$ degrees of freedom. This latter case requires at least 289 electrodes (using the tightest known bound for sampling on a sphere [25]), which is more than that commonly used in most modern systems. Code available at [24].

We now compute the spatial bandwidth required to capture 99% of the signals energy. To that end, we first compute $E_{l_0}^\infty$ for each l_0 .

$$E_{l_0}^\infty = \sum_{l=l_0}^{\infty} \sum_{m=-l}^l |c_{l,m}^{(4)}|^2 = \sum_{l=l_0}^{\infty} \sum_{m=-l}^l |c_{l,m}^{(1)}|^2 |H_{l,m}^{(4)}|^2. \quad (5)$$

Now, $E_{l_0}^\infty = \sum_{l=l_0}^{\infty} \sum_{m=-l}^l \mathbb{E} [|c_{l,m}^{(4)}|^2] = \sum_{l=l_0}^{\infty} \sum_{m=-l}^l \mathbb{E} [|c_{l,m}^{(1)}|^2] |H_{l,m}^{(4)}|^2$. Assuming a spatially white source on the surface of the brain as input to the transfer function, $\mathbb{E} [|c_{l,m}^{(1)}|^2] = 1$, and thus, $E_{l_0}^\infty = \sum_{l=l_0}^{\infty} \sum_{m=-l}^l |H_{l,m}^{(4)}|^2$. Normalizing, the fraction of energy in spherical harmonics $l \geq l_0$ for statistically white inputs passed through spatial filter $H^{(4)}(l, m)$ is:

$$\frac{E_{l_0}^\infty}{E_0^\infty} = \frac{\sum_{l=l_0}^{\infty} \sum_{m=-l}^l |H_{l,m}^{(4)}|^2}{\sum_{l=0}^{\infty} \sum_{m=-l}^l |H_{l,m}^{(4)}|^2}. \quad (6)$$

In Fig. 6, we plot $1 - \frac{E_{l_0}^\infty}{E_0^\infty} = \frac{E_0^{l_0-1}}{E_0^\infty}$, where $E_0^{l_0-1}$ is the energy in the spherical harmonics $l = 0, 1, \dots, l_0 - 1$. The figure shows that $l_0 = 19$ is needed to capture 99% of the signal's energy for the “worst-case” conductivity parameters ($\sigma_3/\sigma_1 = 1/15$). This requires

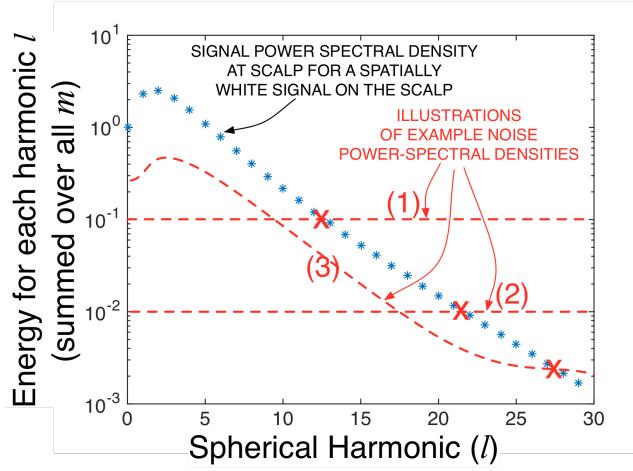


Fig. 7. The plot shows the total energy curve in Fig. 5 on log-scale, illustrating the conceptual shortcoming of the spatial Nyquist-rate analysis: for the purpose of source localization, it does not matter when the transfer function or the average energy becomes small. What is more important is when the energy curve meets the noise level. Three possible noise power-spectral-density (PSD) curves are illustrated. Curves (1) and (2) model (white) sensor noise, while curve (3) models noise from a source within the brain, and hence has a noise PSD similar to the signals, but lower. The points 'X' denote where the signal energy curves meets the noise PSD for each of these, illustrating that with "brain-noise" as the dominant noise source, the limiting number of sensors at which improvements in localization accuracy saturate could be quite high. Code available at [24].

at least $l_0^2 = 361$ sensors. This estimate uses recent work of Khalid, Kennedy, and McEwen [25], which improves on the earlier $2l_0^2$ upper bound [26], and characterizes the required number of sensors to within an approximation by providing an upper bound that matches the lower bound (l_0^2) for signals bandlimited to l_0 bandwidth for moderately small bandwidths ($l_0 < 100$). Similarly, for conductivity parameters $\sigma_3/\sigma_1 = 1/30$, a bandwidth of $l_0 = 17$ is needed that requires at least $l_0^2 = 289$ sensors. The code used to generate these plots is available online at [24].

Nyquist-rate-type lower bounds are usually conservative: temporal sampling is often carried out at rates 1.5-2x the Nyquist rate, and one would expect that to reduce aliasing, a similar rate is required here, increasing the bound for $\sigma_3/\sigma_1 = 1/30$ to 578 which is beyond 512 (the maximum number of sensors used currently). Thus we conclude that even though the estimates on Nyquist rates obtained in [2], [7] underestimate the required number of sensors, currently available (if rarely used) high-density systems are close, and doubling the current number of sensors should reconstruct the scalp EEG within a small error.

We note here the differences in our analysis and the analysis of Srinivasan et al. [2], [7]: (i) Authors in [2], [7] do not add the energy over $m = -l, \dots, l$ to compute the energy for each l , which is one of the errors in their Nyquist rate calculation. As shown in Fig. 5, the average energy curve is very different from the transfer function curve; (ii) The expression for computing the transfer function in [2, Appendix G] is corrected, and presented in a generalized N -sphere format; and (iii) The skull's conductivity is updated with new estimates [22]. With these corrections and updates, the Nyquist rate is observed to be significantly above $l_0 < 10$ observed in [2], [7], which led them to suggest that 128 sensors achieve the Nyquist rate of EEG.

Observe that the number of sensors required to reconstruct scalp EEG faithfully is a lower bound on the required number of sensors to do any reasonable whole-brain imaging: if one can faithfully recover the source of neural activity, one can faithfully recover the EEG signal at any point on the scalp by using standard forward models [11]. However, the converse is *not* true, and thus this lower bound is loose when the goal is source-localization (*i.e.*, imaging). As we discuss in Sections III-B and III-C, Nyquist rate limits severely underestimate the required number of sensors for source localization.

B. Can Nyquist-rate provide a fundamental limit on the required number of sensors for source-localization?

Our main argument is illustrated in Fig. 7. The Nyquist-rate analysis effectively provides a limit on the number of sensors needed to reconstruct the EEG signal with small mean-square error. However, error in reconstructing the EEG signal itself is not a good metric to evaluate goodness of an EEG signal: it is important to understand the goal that the EEG signal is being used for. For instance, for source-localization, where the goal is to image the brain activity, accuracy of source localization is of importance, not the accuracy of EEG signal reconstruction.

In Fig. 8, using simulations and standard algorithms, we estimate improvements in accuracy as number of sensors increase from currently used few hundred sensors to a few thousand sensors. Our main conclusion is that there is significant improvement in resolution at all depths as the number of electrodes increases. Thus, the Nyquist rate analysis likely provides only a very loose lower bound. The reconstruction algorithm used in Fig. 8 reconstructs sources using elastic net [27], which minimizes a weighted sum of an L_2 and an L_1 cost on the source and the squared-error in measured and predicted EEG signals at the sensor locations. The actual source (the "ground truth") is a single dipole located at varying depths (the depth is shown along the x -axis). The reconstruction is observed as a spread of multiple dipole sources. This spread of dipoles is called the "point-spread function" (PSF) [28], which provides a measure of the accuracy of reconstructing a dipole source: the higher the spatial frequencies that are reconstructed, the less spread out would be the PSF. The PSF-width is defined as follows: all points in the reconstruction which are at distances greater

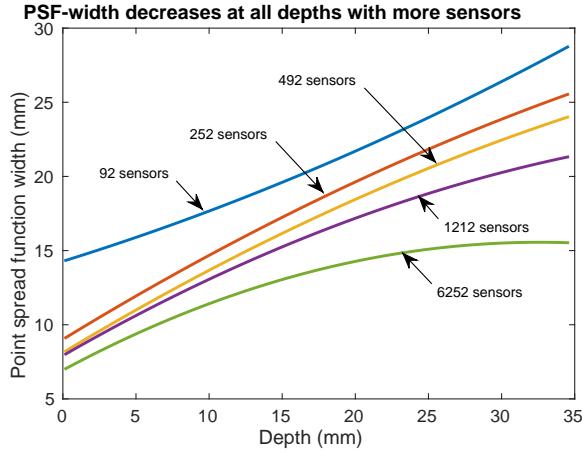


Fig. 8. The width of the “point-spread-function” (PSF) is used as a measure of accuracy of reconstruction. The PSF-width is a measure of the spread in the reconstruction of the source (which is chosen here to be a single dipole). High PSF-widths can be viewed as poor resolution reconstructions. It is plotted here vs depth from the brain-surface for different numbers of sensors for noiseless measurements. There is significant improvement in resolution as the number of electrodes increases. Code available at [24].

than the PSF-width from the true source have PSF-amplitude less than $1/e$ times the max PSF-amplitude. The figure shows that the PSF-width reduces as the number of sensors is increased to even a few thousand. This is because a dipole source contains all spatial frequencies, and an increase in number of sensors increases the spatial frequencies sampled, reconstructing the source with increased accuracy.

It is worth noting that both Nyquist-rate analysis and our simulations in Fig. 8 are noiseless. The number of sensors required is only expected to increase when realistic noise sources are accounted for in order to overcome loss of accuracy due to noise. Nevertheless, as a quick estimate, it is clearly beneficial to increase the number of sensors to be able to sample the frequency at which the signal power is superseded by noise power (see Fig. 7).

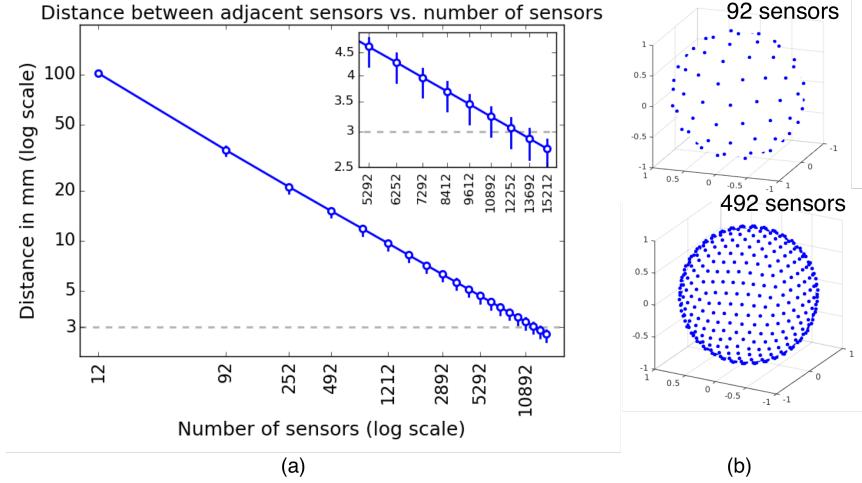


Fig. 9. (a) shows inter-sensor distance vs number of electrodes for electrodes arranged using an icosahedron-based method [29]. (b) shows example icosahedrons for 92 and 492 electrodes to illustrate the arrangement. Code available at [24]. The distances between nearest electrodes for a particular number of electrodes are not perfectly uniform, and the (relatively small) variation in distances is plotted as well (zoomed in for more clarity).

Is it feasible to increase the number of electrodes to such large values? The engineering issues associated with circuit volume and energy/power for sensing circuitry are discussed in Section IV. Our focus here is on the electrode design. How small can the size of the electrode-skin contact be? Are there any signal losses associated with reduced electrode sizes?

To understand this, in Fig. 9 (a), we plot the number of electrodes vs inter-electrode distance for electrodes arranged according to the “icosahedron-based” method developed in [29]. Icosahedron-based constructions (shown in Fig. 9 (b)) are frequently used in various fields such as astronomy, geophysics, and neuroscience [30] to discretize spheres, and satisfy many good covering properties [29]. The figure shows that 10,000 electrodes, the required inter-sensor separation is approximately 3 mm. Is 3 mm inter-electrode distance feasible? Indeed, there are commercially available EEG systems today that are used for rodents that have 3-5 mm inter-electrode spacing, with 1 mm diameter electrodes (see, e.g. [31]). One must remember that electrodes themselves act as (anti-aliasing) spatial filters: electrodes of smaller diameter have higher spatial bandwidths because of their narrow spatial extent. Thus, smaller electrodes allow more signal but also more noise to enter the system, and therefore there is a tradeoff between increasing the number of electrodes to improve degrees of freedom and introducing more noise as a consequence of using smaller electrodes. Another limitation posed by small electrodes is in use of conducting gel commonly used in EEG systems for reduced contact impedance: even if a fast method for

gel-delivery can be discovered, if the gel spreads then it can easily short-circuit electrodes, losing spatial resolution. One may have to use EEG-gel paste that is thicker, or rely on dry electrodes [32], the tradeoffs of which need to be studied further. In summary, proper electrode sizing and design and gel choice is important to sense high-resolution signals, and thus their design is also a parameter in the optimization of the overall EEG system.

C. Fundamental limits on source-localization accuracy of radially-oriented dipoles

We now provide fundamental limits on the accuracy of EEG-based source localization for a given spatial power spectrum of noise. This section advances on our recent work [33]. Why are fundamental limits on source-localization accuracy needed? First, just as in communication theory, converse results and outer bounds here can help inform the design of algorithms that outperform known algorithms. Second, such lower bounds are extremely important to set benchmarks against which algorithms can be compared. Finally, fundamental limits can help abstract and study different aspects of the problem in isolation, which can lead to a systematic approach for algorithm design, circuit design, sensor placement, etc. much like the systematic approach provided by communication theory, which has helped revolutionize practical systems.

The main advancement in our bounds as compared to the existing work of Mosher et al. [5] is that while their bounds hold only for unbiased estimators (since they rely on Cramer-Rao bounds), our bounds apply to all source-localization algorithms. This is particularly important in EEG source localization where commonly used algorithms, such as the Minimum Norm Estimate (MNE) [11], have a bias towards the brain surface [12].

Following the development of source-localization literature [28], [34] (including the fundamental limits in [5]), we restrict our attention to the problem of recovering the location of a single dipole and leave the problem of recovering more general sources to future work. We will provide lower bounds on the average Euclidean distortion $D = \mathbb{E} [\|\mathbf{X} - \hat{\mathbf{X}}\|_2^2] = \sum_{i=1}^3 \mathbb{E} [(X_i - \hat{X}_i)^2]$ in localizing a dipole source in presence of spherically symmetric noise, where $\mathbf{X} = [X_1, X_2, X_3]$ is the location of the dipole in \mathbb{R}^3 . These bounds are then explored numerically for white noise as well as spatially correlated noise arising from a “noise” source within the brain.

The outline of our derivation is as follows: we first compute the power spectrum generated by a dipole at the input of the brain-to-scalp channel (characterized in Theorem 1). We then bound the capacity of this channel under an input power constraint specified by the dipole’s power spectrum. Finally, we use a distortion-rate function of the dipole source calculation applied to the channel capacity to obtain a fundamental limit on the source-localization accuracy. The information-theoretic nature of our derivation ensures that the bound is applicable to all source-localization algorithms, and not just the unbiased ones, and thus goes beyond the prior work in [5] where Cramer-Rao-type bounds are obtained that are only applicable to algorithms that are unbiased estimators.

For simplicity, as a first step, we further restrict our attention to radially-oriented dipoles, which are easier to analyze because their potentials induce fields with azimuthal symmetry [2] when z-axis is aligned with the dipole orientation (the axes can then be rotated to obtain the field generated by a dipole not on the z-axis). The dipole is assumed to be located at radius r_z from the center of the sphere (unknown at the receiver), have intensity I_d , and (a small) distance between poles d_z .

Lemma 2 (Scalp potential generated by a single dipole): For a single radially-oriented dipole located on the z-axis ($\theta = 0$) at radius r_z in the first sphere of the 4-sphere model with current I and pole separation d_z , the potential at any point for $r_z < r \leq r_1$ is given by:

$$\Phi^{(1)}(r, \theta) = \frac{Id_z}{4\pi\sigma_1 r_z^2} \sum_{l=1}^{\infty} \left[\alpha_l \left(\frac{r}{r_1} \right)^l + \left(\frac{r_z}{r} \right)^{l+1} \right] l P_l(\cos \theta). \quad (7)$$

Here $P_l(\cdot)$ ’s are the Legendre polynomials [21], $\alpha_l = \frac{1}{\gamma_l^{(1)}} \left(\frac{r_1}{r_z} \right)^{l+1}$, and $\gamma_l^{(1)}$ is given in Theorem 1. The coefficients $A_{l,m}^{(1)}$ and $B_{l,m}^{(1)}$ in (2) are given by $B_{l,0}^{(1)} = \frac{Id_z}{4\pi\sigma_1 r_z^2} l \sqrt{\frac{2}{2l+1}} \left(\frac{r_z}{r_1} \right)^{l+1}$, $A_{l,0}^{(1)} = B_{l,0}^{(1)} / \gamma_l^{(1)}$, and $A_{l,m}^{(1)} = B_{l,m}^{(1)} = 0$ for any $m \neq 0$. The “power” content of the scalp potential $p_l(I, d_z, r_z^2)$ in l -th spherical harmonic is given by $p_l(I, d_z, r_z^2) = (A_{l,0}^{(4)} + B_{l,0}^{(4)})^2$, which can be obtained using Theorem 1 to compute $A_{l,0}^{(4)}$ and $B_{l,0}^{(4)}$.

Proof: Equation (7) for the potential produced by the dipole within the innermost sphere is taken directly from [2, Pg. 562, (G.1.9)]. All that remains is to compute $B_{l,0}^{(1)}$ (the rest follow for spherical symmetry and from Theorem 1). This is done by comparing (2) and (7) to obtain $B_{l,m}^{(1)}$, keeping in mind that $\int_{\theta=0}^{\pi} (P_l(\cos \theta))^2 \sin \theta d\theta = \frac{2l+1}{2}$. ■

We will use the capacity of the dipole-to-scalp channel to obtain these fundamental limits. To compute this capacity, we will use the transfer function from the dipole to the scalp obtained in Lemma 2.

The bounds are given by the following theorem.

Theorem 2 (Source-localization fundamental limits): For the 4-sphere model described in Section II, for a radially oriented dipole and a spatial noise spectrum of power N_l in the (l, m) -th spherical harmonic (assumed to be spherically symmetric, *i.e.*, equal for all m for a given l), a lower bound on the average localization distortion $D = \mathbb{E} [\|\mathbf{X} - \hat{\mathbf{X}}\|_2^2]$ for a dipole uniformly distributed inside the innermost sphere is given by:

$$D \geq \frac{3}{2\pi e} V_1^{\frac{2}{3}} 2^{-2\bar{C}/3}, \quad (8)$$

where $V_1 = \frac{4}{3}\pi r_1^3$ is the volume of the innermost sphere, $\overline{C} = \sum_{l=0}^{\infty} \frac{(2l+1)}{2} \log_2 \left(1 + \frac{p_l(I, d_z, r_z^2)}{(2l+1)N_l} \right)$ is an upper bound on the capacity of the dipole-to-scalp channel, and p_l is the power of the potential generated by the dipole on the scalp surface given by Lemma 2.

Proof: Because of azimuthal symmetry, the terms with $m \neq 0$ in the summation can be ignored, and one can simply compute power p_l allocated to the l -th spatial frequency (and noise in this frequency). A rotation of the axes will introduce power into spherical harmonics with $m \neq 0$, but the overall power in the l -th harmonic (summed over all m) will remain the same. Thus, this power depends on the parameters r_z , I , and d_z , but not on the θ and ϕ coordinates of the dipole.

For analytical simplicity, we first use the Shannon lower bound [35, Pg. 338] to obtain a closed-form expression for the distortion-rate function for reconstructing the source. The mutual information between the location \mathbf{X} and its reconstruction $\widehat{\mathbf{X}}$ can be bounded as follows:

$$I(\mathbf{X}; \widehat{\mathbf{X}}) = h(\mathbf{X}) - h(\mathbf{X} | \widehat{\mathbf{X}}). \quad (9)$$

Analyzing the first term in (9),

$$h(\mathbf{X}) = - \int_{x \in \mathcal{S}_1} \frac{1}{V_1} \log_2 \frac{1}{V_1} dx = \log_2 V_1, \quad (10)$$

where \mathcal{S}_1 is the innermost sphere in which the dipole lies, and V_1 is the volume of the sphere.

Analyzing the second term in (9),

$$\begin{aligned} h(\mathbf{X} | \widehat{\mathbf{X}}) &= h(X_1, X_2, X_3 | \widehat{X}_1, \widehat{X}_2, \widehat{X}_3) = h(X_1 - \widehat{X}_1, X_2 - \widehat{X}_2, X_3 - \widehat{X}_3 | \widehat{X}_1, \widehat{X}_2, \widehat{X}_3) \\ &\leq h(X_1 - \widehat{X}_1, X_2 - \widehat{X}_2, X_3 - \widehat{X}_3) \leq h(X_1 - \widehat{X}_1) + h(X_2 - \widehat{X}_2) + h(X_3 - \widehat{X}_3) \\ &\stackrel{(a)}{\leq} \frac{1}{2} \log 2\pi e D_1 + \frac{1}{2} \log 2\pi e D_2 + \frac{1}{2} \log 2\pi e D_3 = \frac{3}{2} \log 2\pi e \sqrt[3]{D_1 D_2 D_3} \\ &\stackrel{(b)}{\leq} \frac{3}{2} \log 2\pi e \frac{D_1 + D_2 + D_3}{3} = \frac{3}{2} \log 2\pi e \frac{D}{3}, \end{aligned} \quad (11)$$

where $D_i = \mathbb{E} \left[(X_i - \widehat{X}_i)^2 \right]$, (a) follows from the fact that Gaussian distribution maximizes the differential entropy for a given second moment of the random variable, and (b) follows from the fact that arithmetic mean of positive real numbers dominates their geometric mean.

From (9), (10), and (11), we get:

$$I(\mathbf{X}; \widehat{\mathbf{X}}) \geq \log_2 V_1 - \frac{3}{2} \log 2\pi e \frac{D}{3}. \quad (12)$$

We now show that this mutual information can be upper bounded by the channel capacity (for power allocation across different l as generated by a single dipole on the scalp surface):

$$I(\mathbf{X}; \widehat{\mathbf{X}}) \leq C_{\text{dipole-to-scalp}} \stackrel{(a)}{\leq} \sum_{l=0}^{\infty} \frac{(2l+1)}{2} \log_2 \left(1 + \frac{p_l(I, d_z, r_z^2)}{(2l+1)N_l} \right) =: \overline{C}, \quad (13)$$

where N_l is the scalp noise in the (l, m) -th spatial frequency, and p_l is the power in the l -th frequency (summed up over all $m = -l, -l+1, \dots, l-1, l$), at the scalp generated by a single dipole, which is obtained in Lemma 2. The RHS of (13), which is an upper bound on $C_{\text{dipole-to-scalp}}$, is denoted by \overline{C} . How do we obtain this upper bound (a)? We allow the dipole to split its power p_l freely across $m \in \{-l, \dots, l\}$. Thus, the dipole has available $(2l+1)$ identical parallel channels for each l (identical because the noise variance is the same in each channel) with a sum power constraint of p_l . For maximizing the capacity, it will distribute the power p_l for l -th harmonic equally across all $(2l+1)$ channels ($m = -l, -l+1, \dots, l-1, l$). The theorem now follows from (12) and (13). ■

Remark 1: In the proof, we allow the dipole to encode any information it wants to through its dipole moment by relaxing the equality constraint on its intensity to an average intensity constraint. This is an optimistic assumption, and one reason why the obtained bounds are loose. Improvements on the bounds as well as comparison with upper bounds are required to understand the accuracy of this approach.

Remark 2: Note that this is a bound on capacity with a *single* time measurement but with infinite scalp electrodes. The capacity of the channel depends on I, d_z and r_z . A lower available power, e.g. because of small r_z (which leads to a larger decay), naturally increases the reconstruction error. Often, the experimentalist can reduce source-localization error by repeating the experiment over multiple trials (and our bounds can be easily extended to that case). However, in some cases trials can be very limited, e.g. in localizing “focal point” of an epileptic seizure where the required number of EEG electrodes (and whether EEG constitutes a reliable imaging technology) is a hotly debated issue [36].

1) Numerical results: White noise at sensors: We first consider the case when only sensor noise is present which is assumed to be white (*i.e.*, spatially uncorrelated) and Gaussian. This models the noise present in circuitry and electrodes in the system. Thus, $N_l = \sigma_z^2$ for all l . Note that this means that the noise for higher l (added over all m) is higher, which is the appropriate definition of white noise in spherical harmonics (e.g. see [37]). From the numerical results shown in Fig. 10, reducing sensor noise (even by 3 dB) can have significant impact on improving the source localization accuracy. Thus, improving circuit designs to lower amplifier

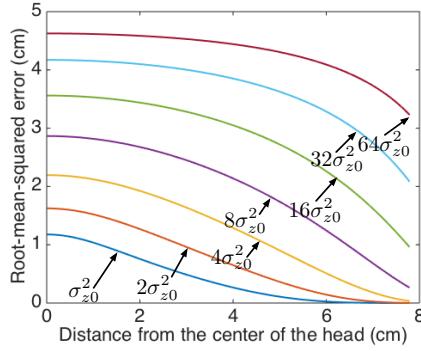


Fig. 10. A lower bound on the tradeoff between \sqrt{D} (the root-mean-squared error in cm) and the depth of the source for different white-noise levels for the case when only white sensor noise is present. The error increases with depth, as expected. As this sensor noise is lowered (the curves are separated by 3 dB, *i.e.*, a factor of 2), the lower bound on the error at different depths reduces substantially, suggesting that there are significant gains to reducing the sensing noise. The parameter values are $\sigma_{z0}^2 = 10^{-5}$ and $\frac{Id_z}{4\pi\sigma_1^2} = 1$.

and ADC noise as well as improved electrode design to have low contact impedance can help improve source-localization accuracy significantly when sensor noise is the dominant noise source.

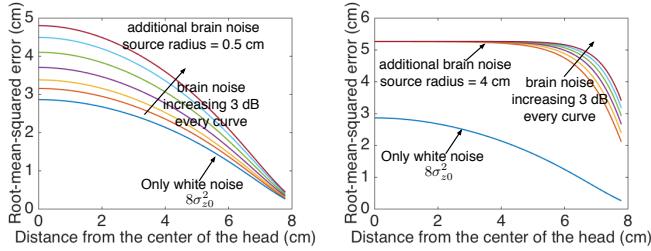


Fig. 11. The lower bound on root-mean-squared error (in cm) as the depth of the source and the noise level of a deep source (that generates a white signal at radius 0.5 cm from the sphere center in the left figure, and at 4 cm in the right figure, and hence a correlated signal at the scalp surface). The white noise present is $8\sigma_0^2$ where σ_0^2 is equal to that chosen in Fig. 10. The lowest (dark blue) curve in both plots is the plot with only white noise, and thus matches with the $8\sigma_{z0}^2$ plot in Fig. 10. For the rest, an additional brain source noise that increases by a factor of 2 (3 dB) in each plot is used, starting from a noise power that is half the power of the dipole.

2) *Numerical results: Noise due to a brain source:* What quantities contribute to noise in EEG signals? Beyond thermal noise and the noise in the sensing circuitry, this includes any artifact noise that has not been rejected or canceled out (such as muscle and eye artifacts, “EMG” and “EOG” respectively), any in-band electromagnetic radiation from outside the brain, and also “noise” from within the brain. This last source of noise, namely, signals generated by brain activity not relevant to the neuroscientific goal, are of interest here. In particular, we are interested in noise generated by sources that are uncorrelated to the neural activity that the experimentalist wants to study⁸.

Some of the above noise terms can be reduced, *e.g.*, the sensing circuitry noise can be lowered by improved amplifier and ADC design [37], the artifact noise is (at least partially) regressed out, and in-band radiation is shielded against by performing the experiment in Faraday cages. The “brain noise” term is probably the most difficult to reduce because estimating it requires solving another inverse problem.

Here we study the effects of noise arising from within the brain under simplifying assumptions of noise isometry (*i.e.*, the noise distribution is spherically symmetric). To model this, we assume that a spatially white source is producing noise at a certain depth ($r = 0.5$ cm and 4 cm in Fig. 11) inside the brain. As the signals propagate outside, they get correlated because of the low-pass filtering effect (Theorem 1), and the eventual addition to EEG signal is highly correlated. The resulting effect on source-localization accuracy is qualitatively different from that of white noise, and is illustrated in Fig. 11. The figure shows that a deeper noise source (at $r = 0.5$ cm) has a smaller effect on source-localization accuracy of shallow sources. This is because a deeper noise source suffers more severe attenuation, especially in high spatial frequencies, thus adding substantially smaller noise in frequencies that carry high resolution information. Interestingly, the accuracy near the brain surface (*i.e.*, close to $r_1 = 7.8$ cm) is not affected substantially by the presence of this deep noise source even as its amplitude is varied significantly. On the other hand, even a low-amplitude shallow noise source (*e.g.* at $r = 4$ cm) can result in substantial errors, especially in localizing deep sources of interest.

IV. INFORMATION-THEORETIC LENS ON EEG ENGINEERING

Fundamentally, high-density neural sensing requires low-power and low-volume sampling, communication, and inference. Naturally, one expects tools from information theory to prove useful in designing these systems. For instance, communicating the sensed EEG

⁸If there is correlation between the source of interest and another source, we believe the problem should be treated as a two-source problem.

data raises interesting communication challenges. The raw data rates can be quite high for the high-density systems proposed here (tens to hundreds of megabits per second, or more), and thus require substantial energy sensing and communication energy (tens of Watts or more), reducing which is an interesting and important challenge. Techniques tied to the goal of acquiring EEG signal can be used to reduce energy consumption. E.g., when monitoring an epilepsy patient, the end user might only need high resolution data at the onset of the seizure. A local processing can detect seizure and communicate stored data only on positive detection. In other cases, the experimentalist might only be interested in executing a machine-learning algorithm (e.g., performing a classification or regression on the sensed data that are frequently used in brain-machine interfaces [38], [39] and in many neuroscientific analyses). In such cases, not all the sensors sense may relevant data. Indeed, most of the sensors (in particular those that are sensing data that is less relevant to the end goal) may only need be sensed and communicated at a low-resolution, and only some sensors might need a high-resolution representation. This allows for drastic compression by exploiting information-theoretic techniques in rate and energy-constrained learning and inference that quantify the relevance of each sensors data by estimating the system parameters [40]–[44].

However, when the problem is one of communicating data for imaging, or for an inference analysis that is not known in advance, it can be the case that all sensors need to be sensed at a high-resolution. In such situations, the required power can easily exceed tens of watts for high-density EEGs discussed here (as discussed in Section I). Can information-theoretic techniques be exploited to reduce power in these situations? As shown in Theorem 1, the brain surface to scalp transfer function acts as a spatial low-pass filter. The resulting spatial power-spectral-density of EEG as it passes through the CSF, the skull, and the scalp (in a 4-sphere model) is shown in Fig. 5. Because most energy is concentrated in lower frequencies, the induced correlation is quite high (see Fig. 13). Further, noise due to EMG and EOG (muscle and eye artifacts) is also spatially highly correlated (as has been observed experimentally, e.g. [13]). Thus the measured signal indeed has high correlations. Can this correlation be harnessed to reduce the energy and volume requirements on the sensing circuitry? Note the interesting twist here: while it is known that the decay of high spatial frequencies lowers EEG’s resolution (as can also be seen through Theorem 2), it might be possible to exploit the spatial correlations induced by this decay to reduced sensing energy and volume!

One might be tempted to ask: if the sensors are sensing highly correlated data, do we need to have high-density sensing in the first place? Indeed, as noted earlier (and illustrated in Fig. 1), there is important information buried even in the “less significant” bits of each observation. Because of decay of high spatial frequencies, these lower order bits are the ones carrying high-resolution (high spatial frequency) information. They capture subtle sources that generate small variations in potential due to a combination of their shallow depth or small spatial extent. Perhaps as an acknowledgement of this fact, many state-of-the-art systems record each electrode at a very high precision (e.g. 24 bit ADCs [17]).

A. Referencing strategies for low-noise sensing of EEG signals

Potentials are relative: the potential at any point is always measured with respect to another point. Ideally, EEG signals on the scalp should be referenced with respect to infinity. In practice, they are referenced with respect to potential at some point on the body, such as an ear or a shoulder. For source-localization and imaging, it is useful to have the potential be estimated with respect to a global reference. Thus, the goal in the mechanisms in this section is to report signal values at all electrodes with respect to a global reference even if signals are actually sensed differentially and thus may require a computational step post sensing. For implementation simplicity, we do not allow using sophisticated information-theoretic binning-techniques (e.g. [45]). For instance, one might sense signals differentially, *i.e.*, measure an electrode’s potential with respect to a nearby electrode, reducing the number of bits to sense because of the high correlation. The question of interest here is what differential sensing architectures, *i.e.*, referencing mechanisms, reduce noise aggregation at the post-processing computation step.

1) *Referencing mechanisms as trees*: We first present a simple observation: every referencing mechanism can be represented as an associated tree. Each electrode corresponds to one node in this tree. The global reference electrode is the root node. All nodes referenced directly against the global reference are the nodes at level 1. All nodes referenced directly against level 1 nodes are level 2 nodes, and so on. In the tree, a node’s parent is the node the corresponding electrode is referenced against. Because each electrode is referenced against exactly one electrode, each node has exactly one parent, and thus the corresponding graph for any referencing strategy is a tree.

Let us consider two simple strategies to illustrate this association. In the first case, shown in Fig. 12 (left), is a commonly used system where all electrodes are referenced directly against a global reference (sometimes called “unipolar” or “direct global referencing”). The corresponding tree, also shown in Fig. 12 (left), has one root node with all other nodes as its children. In the second case, that we will call “sequential bipolar referencing” (or simply sequential referencing) each electrode is referenced to an electrode prior to it in a chain of electrodes (see Fig. 12, middle) leading to a tree where each node has 1 child, which is related to the strategy of communicating innovations analyzed by Kim and Berger [18], [19]. We now discuss another possible mechanism: hierarchical referencing.

2) *The hierarchical referencing mechanism, and comparison with sequential bipolar and direct global referencing*: The mechanism proposed in this work is illustrated in Fig. 12 (right). Instead of using sequential bipolar referencing, we arrange the electrodes in a tree pattern with number of levels that can be optimized based on spatial correlations. The mechanism is easy to design. For instance, for placement of electrodes along a grid, one can simply think of the grid as a hierarchical subdivision of squares. In case of placement of electrodes on a spherical surface, commonly used techniques are inherently recursive (e.g., icosahedron bisections used in EEG grids) and lend themselves naturally to hierarchical constructions of unequal but roughly constant degree per level. We now discuss why this strategy improves on the sequential bipolar strategy and the direct global referencing (unipolar) strategy.

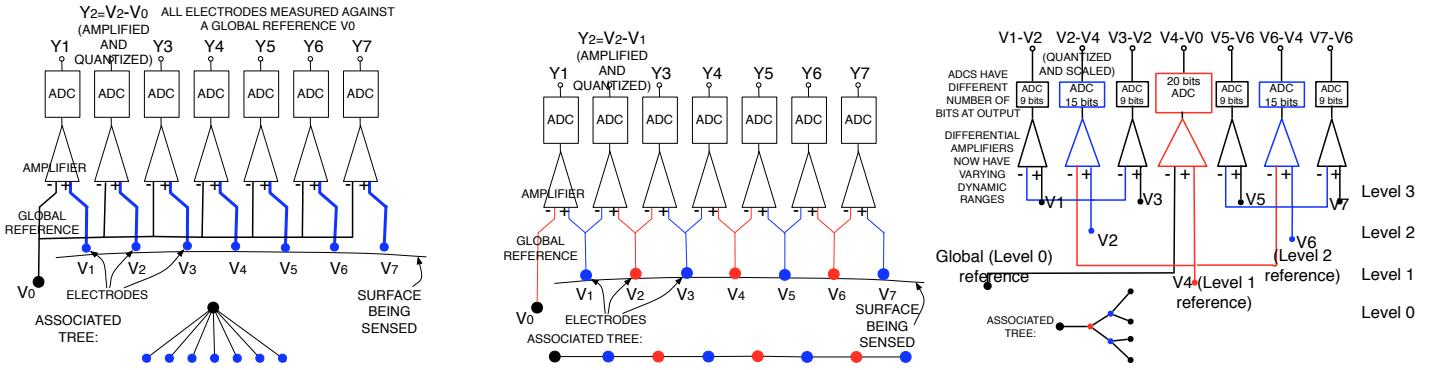


Fig. 12. (left) Direct global referencing, and the associated tree, where every electrode is referenced directly against a global electrode. (middle) Sequential bipolar referencing and the associated tree. (right) The hierarchical referencing strategy. For clarity, we indicate on top the differences of voltages being measured, and ignore quantization and thermal noise added by the circuitry. If the circuits were noiseless, and actual differences of voltages were being measured, all voltages can be recovered with respect to the global reference in just three steps. For instance, $V_3 - V_0 = (V_3 - V_2) + (V_2 - V_4) + (V_4 - V_0)$. Of course, what are actually measured are $Y_i = V_i - V_{\text{ref}(i)} + Z_i$, where Z_i is the noise term. Thus what is recovered is the sum of these voltages added to a noise whose variance is the sum of the variance of noise at each electrode.

B. Comparison of the three strategies: theory, numerical results, and experimental studies

To demonstrate the theoretical benefits of our strategy, we first consider a nonstationary vector source:

$$\mathbf{X}_{i+1}^m = \mathbf{X}_i^m + \nu_i^m \quad (14)$$

where $\nu_i^m \sim \mathcal{N}(0, \sigma_\nu^2 \mathbb{I})$ is independent of $\mathbf{X}_j^m, j \leq i$. \mathbf{X}_0^m is assumed to be zero. Overall mean-squared distortion (averaged over realizations of ν_i^m) is constrained to be smaller than D for each \mathbf{X}_i^m (*i.e.*, worst-case over sensors). For compressing such sources in a centralized fashion, it is well known that sequential-referencing-type strategies are suboptimal [18], [19]. However, the decentralized nature of the problem posed here rules out strategies such as “Differential Pulse-Coded Modulation” (DPCM, see [20] for an information-theoretic analysis) that is used extensively in modern communication systems because they require the “transmitter” (which in this case are distributed sensors) to know the receiver’s estimate (*i.e.*, to use DPCM here would require a sensor to know the *quantized reconstruction* of signal sampled by another sensor, which is not the case).

For simplicity, we use rate-distortion functions derived for large m . The results in Section IV-B1 – that are numerical results (on idealized models) as well as experimental results – are obtained using realistic scalar quantization.

Global referencing: The rate distortion function is given by $R_i = \frac{1}{2} \log \frac{\sigma_i^2}{D} = \frac{1}{2} \log \frac{i \sigma_\nu^2}{D}$, and thus the total rate required by this strategy is $R_{\text{total}}^{\text{global}} = \sum_{i=1}^n R_i = \Omega\left(n \log \frac{n \sigma_\nu^2}{D}\right)$.

Sequential referencing strategy: In this case, each sensor is sensing ν_i^m , which it quantizes with R_i rate. Then, the average distortion at the i -th sensor is given by $D_i = \sigma_\nu^2 2^{-2R_i}$. Further, the distortions at different sensors are independent (because ν_i^m are independent across sensors). Thus, we obtain a total distortion at the last sensor as the sum of all distortions:

$$\begin{aligned} D &= \sum_{i=1}^n D_i = \sum_{i=1}^n \sigma_\nu^2 2^{-2R_i} = \sigma_\nu^2 \sum_{i=1}^n 2^{-2R_i} = n \sigma_\nu^2 \frac{1}{n} \sum_{i=1}^n 2^{-2R_i} \\ &\stackrel{\text{(Jensen's inequality)}}{\geq} n \sigma_\nu^2 2^{-2 \frac{1}{n} \sum_{i=1}^n R_i}. \end{aligned}$$

Thus,

$$R_{\text{total}}^{\text{seq}} = \sum_{i=1}^n R_i \geq \frac{n}{2} \log \frac{n \sigma_\nu^2}{D} = \Omega\left(n \log \frac{n \sigma_\nu^2}{D}\right)$$

Hierarchical referencing strategy: Let D_l denote the maximum distortion at the l -th level in the tree. Choose⁹ $D_l = \frac{D 2^l}{n}$, where $n = 2^{L+1}$, where L is the total number of levels in the tree. Then, it is easy to check that the total distortion is bounded by D :

$$\sum_{l=1}^{L=\log(n)-1} D_l = \frac{D(2^{L+1} - 1)}{n} = \frac{D(2^{L+1} - 1)}{2^{L+1}} \leq D. \quad (15)$$

⁹How did we arrive at this choice? We simply solved the constrained optimization problem: $\min_{D_i, i=1, \dots, L} \sum_{i=1}^L 2^{i-2} \log_2 \left(\frac{n \sigma_\nu^2}{2^i D_i} \right)$ s.t. $\sum_{i=1}^L D_i \leq D$.

The above analysis assumes that the distortions in the tree at each level are independent, which they are because of the randomness in codebook-choice at each level. The total rate required by this strategy is:

$$\begin{aligned}
R_{\text{total}}^{\text{hier}} &= \sum_{i=j}^n R_j \stackrel{(a)}{=} \sum_{l=1}^L 2^{l-1} \frac{1}{2} \log \left(\frac{n\sigma_{\nu^2}}{2^l} \times \frac{1}{D_l} \right) = \sum_{l=1}^L 2^{l-2} \log \left(\frac{n^2\sigma_{\nu}^2}{2^{2l}D} \right) \\
&\stackrel{l'=L-l+1}{=} \sum_{l'=1}^L 2^{L-l'-1} \log \left(\frac{n^2\sigma_{\nu}^2}{2^{2L-2l'+2}D} \right) = \sum_{l'=1}^L \frac{n}{2^{l'+2}} \log \left(\frac{2^{2(l'-1)}\sigma_{\nu}^2}{D} \right) \\
&= \sum_{l'=1}^L \frac{n}{2^{l'+2}} \left(2(l'-1) + \log \frac{\sigma_{\nu}^2}{D} \right) \leq n \sum_{l'=1}^{\infty} \frac{1}{2^{l'+2}} \left(2(l'-1) + \log \frac{\sigma_{\nu}^2}{D} \right) = O \left(n \log \frac{\sigma_{\nu}^2}{D} \right).
\end{aligned}$$

where the term 2^{l-1} in (a) comes from the number of sensors at the l -th level of the tree, and the term $\frac{n\sigma_{\nu^2}}{2^l}$ is an upper bound on the variance of the difference observed by a node at the l -th level (with respect to an appropriate node at the $(l-1)$ -th level). Thus the required rate using hierarchical referencing grows only linearly with the number of sensors for a given worst-case (over sensors) average distortion, in comparison with $\Omega(n \log n)$ for the alternative strategies. The total bit savings of hierarchical referencing in comparison with global or sequential referencing can thus be arbitrarily large.

1) *A numerical and experimental case study: ultra high density EEG:* In this section, we obtain estimates of power savings using hierarchical referencing for ultra high density EEG sensing (with up to 9331 electrodes; conventional systems use only up to 512 electrodes, typically much smaller). The code for this case study is available online at [24].

Algorithm for optimal tree search: To state the problem of optimal tree-search formally, there are n sensors sensing voltages $V_i, i = 1, \dots, n$. A goal can be to find the referencing tree \mathcal{T} that solves the following problem:

$$\min_{\mathcal{T} \text{ s.t. } E_{\mathcal{T}} \leq E_{\max}} \max_{i=1, \dots, n} \mathbb{E}[(\hat{V}_i - V_i)^2], \quad (16)$$

where V_i is the potential at electrode i with respect to the global reference, \hat{V}_i is its estimate, $E_{\mathcal{T}}$ is the energy required by measurements corresponding to the measurement tree, and E_{\max} is the total available energy. By taking maximum over $i = 1, \dots, n$, we ensure that the maximum distortion over all electrodes is minimized. The expectation in (16) is taken over the (random) values of input V_i .

The search for the optimal referencing tree can be hard: the number of possible trees for n nodes is given by Cayley's formula [46], n^{n-2} , which is super-exponential in n . Therefore, we limit ourselves to obtaining the optimal (symmetric) hierarchical strategy, and comparing it to the global referencing strategy.

Assuming that the spatial correlation between sensors as a function of distance between them is known in advance, a brute-force algorithm provides a solution that is still polynomial time in number of sensors for a symmetric hierarchical tree when the signal is spatially stationary (as is the case in our idealized model for EEG signals).

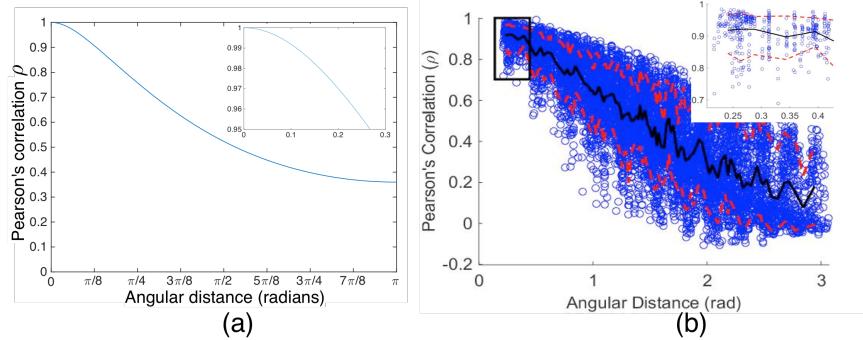


Fig. 13. (a) Theoretically computed EEG spatial correlations between two points separated by θ angular separation on the scalp. The correlations are computed using the 4-sphere model with parameters as described earlier, *assuming a white signal on the surface of the brain*. Correlations estimated using more realistic models such as fibre-tracks will likely only be larger because the signals on the surface of the brain are likely to be more slowly spatially changing than white signals. (b) Experimentally computed correlations using data obtained from a 128 electrode EEG system in Behrmann lab at CMU (taken from [47]). The experimental results in [47] show that up to 3 bits of ADC can be reduced for all ADCs at the highest level in the hierarchical referencing tree (which are most of the ADCs).

Computing correlations: Because of lack of ultra-high-density EEG systems, empirical measurements of correlations at small distances are lacking. Therefore, we use (a) idealized 4-sphere models to obtain theoretical estimates (see Fig. 13); and (b) extrapolation based on current state-of-the-art sensing system of 128 electrodes. The theoretical results are expected to underestimate the correlations somewhat (and therefore, underestimate the energy savings using our strategy) by assuming that the signal at the surface of the brain is white. One expects the generated signal itself to have more content in lower frequencies because of spatial correlation in neuronal activity, especially in the cortex. Nevertheless, our analysis using ADC energy models [37], [47], [48] shows that bit savings can exceed 3 bits on average, and energy savings can exceed 3x, using the hierarchical architecture in both simulations on idealized head-models [37], [48] and extrapolations from experimental data [47] (obtained in the Behrmann lab).

V. DISCUSSIONS AND CONCLUSIONS

This paper explores informational limits to EEG-based noninvasive neural sensing through spatial-Nyquist-rate arguments, simulations, and information-theoretic limits. We observed that for source-localization, the required number of sensors can be significantly larger than what state-of-the-art EEG systems use today. In fact, clinically used EEG systems largely have less than 100 electrodes, and for source localization of any accuracy, clinicians rely on either the bulky and expensive MEG equipment, and/or invasive monitoring using Electrocorticography and use of depth electrodes. We contend that the informational limits of EEG have not been achieved, and thus there is need to explore these limits through a mix of theory (impossibility results as well as achievable strategies), experiments, and advanced system design.

One of the major shortcomings of the work here is the simplifying spherical head models that are used in most of our theory, analyses, and simulations. E.g., real heads impose constraints on where sensors can be placed, on densities of sensors at various locations, and on sources of activity. Further, for improving imaging accuracy for real heads by increasing number of electrodes, having accurate forward models (*i.e.*, the transfer function from the neural source to the scalp) will become very important. It has been noted that spherical head models are not sufficiently accurate approximations even with current low-density systems [49], and with an increased number of sensors, the importance of precise models for each individual head will only increase. This can be done using a combination of complementary imaging techniques such as MRI and electrical impedance tomography (EIT). While MRI scans provide thickness of various layers with an accuracy that neither MEG nor EEG can achieve (~ 1 mm), they do not yield conductivities of these layers. Because conductivities of these layers can vary across individuals [50] (and even within an individual at different points), MRI itself may not yield sufficiently accurate forward models. EIT uses injection of small currents to estimate conductivities, and thus while it has extremely low-resolution imaging reconstructions currently, it can complement MRI by providing estimates of conductivity. These can then be used to obtain forward models by solving Poisson's equation using boundary element techniques. From the perspective of this paper, any remaining uncertainty in the parameters results in a "compound" [51, Ch. 7] brain-to-scalp channel. This view could pave the way for further information-theoretic analyses, e.g., by extending the approach used here to obtain fundamental limits, and developing universal source-localization strategies that provide guaranteed reconstructions despite uncertainty in channel parameters.

A missing piece in the current results is obtaining fundamental limits on source-localization accuracy as the number of sensors scale. Our current results provide an asymptotic (infinite-sensor) analysis, and understanding the effect of a finite number of sensors would require analyzing the effects of aliasing on the transfer function. While it is possible to engineer electrodes that reduce aliasing (because electrodes act as anti-aliasing spatial filters), it is unclear if eliminating aliasing is indeed optimal.

Beyond source-localization, an exciting problem of current interest is estimating functional connectivity of brain networks, and obtaining directions of information flows [52], [53], and thus it is also of interest to understand fundamental limits on required number of sensors for connectivity-estimation. The dual to this problem, which utilizes an understanding of anatomical connectivity to improve source-localization accuracy [54] is a problem that has concrete priors on the activity's evolution, and should be amenable to similar analyses.

On the implementation side, to make an ultra-high-density EEG system portable will likely require reducing measurement energy below that achieved by the hierarchical referencing strategy proposed here. We believe that techniques that combine sensing, communication, and inference (e.g. [40]–[44]) will prove extremely relevant in this context. Finally, while we provide an achievable strategy for the problem of distributed lossy compression considered here that outperforms conventional strategies in both theoretical and experimental analyses, it is unclear if it is close to optimal. Developing converse techniques for this distributed sensing setup is an exciting new information-theoretic problem that falls out of this study.

ACKNOWLEDGMENTS

We thank Ted Huppert and Jay Deep Sau for helpful discussions, and Marlene Behrman, Lori Holt, Haewon Jeong, Shawn Kelly, Mike Tarr, and Jeffrey Weldon for motivating these problems from an interdisciplinary neuroscientific and engineering viewpoint. We also thank the CMU BrainHUB program for facilitating discussions between engineers and neuroscientists that led to the problems formulated and addressed here. We acknowledge support from NSF CCF-1350314, CMU BrainHUB, and by Systems on Nanoscale Information fabriCs (SONIC), one of the six SRC STARnet Centers, sponsored by MARCO and DARPA. Praveen Venkatesh was in part supported by a Henry L. Hillman Presidential Fellowship.

REFERENCES

- [1] S. P. Ahlfors, J. Han, J. W. Belliveau, and M. S. Hämäläinen, "Sensitivity of MEG and EEG to source orientation," *Brain topography*, vol. 23, no. 3, pp. 227–232, 2010.
- [2] P. L. Nunez and R. Srinivasan, *Electric fields of the brain: the neurophysics of EEG*. Oxford university press, 2006.
- [3] S. Klamer, A. Elshahabi, H. Lerche, C. Braun, M. Erb, K. Scheffler, and N. K. Focke, "Differences Between MEG and High-Density EEG Source Localizations Using a Distributed Source Model in Comparison to fMRI," *Brain topography*, vol. 28, no. 1, pp. 87–94, 2015.
- [4] S. J. Luck, *An introduction to the event-related potential technique*. MIT press, 2014.
- [5] J. C. Mosher, M. E. Spencer, R. M. Leahy, and P. S. Lewis, "Error bounds for EEG and MEG dipole source localization," *Electroencephalography and clinical Neurophysiology*, vol. 86, no. 5, pp. 303–321, 1993.
- [6] A. H. Marblestone, B. M. Zamft, Y. G. Maguire, M. G. Shapiro, T. R. Cybulska, J. I. Glaser, D. Amodei, P. B. Stranges, R. Kalhor, D. A. Dalrymple, D. Seo, E. Alon, M. M. Mahabiriz, J. M. Carmena, J. M. Rabaey, E. S. Boyden, G. M. Church, and K. P. Kording, "Physical principles for scalable neural recording," *Frontiers in Computational Neuroscience*, vol. 7, 2013.
- [7] R. Srinivasan, D. M. Tucker, and M. Murias, "Estimating the spatial Nyquist of the human EEG," *Behavior Research Methods, Instruments, & Computers*, vol. 30, no. 1, pp. 8–19, 1998.

- [8] C. R. Rao, "Information and the accuracy attainable in the estimation of statistical parameters," *Bulletin of the Calcutta Mathematical Society*, vol. 37, pp. 81–89, 1945.
- [9] P. Stoica and R. L. Moses, "On biased estimators and the unbiased Cramér-Rao lower bound," *Signal Processing*, vol. 21, no. 4, pp. 349–350, 1990.
- [10] M. Hardy, "An illuminating counterexample," *American Mathematical Monthly*, pp. 234–238, 2003.
- [11] R. R. Ramrez, "Source localization," *Scholarpedia*, vol. 3, no. 11, p. 1733, 2008. [Online]. Available: http://www.scholarpedia.org/article/Source_localization
- [12] F.-H. Lin, T. Witzel, S. P. Ahlfors, S. M. Stufflebeam, J. W. Belliveau, and M. S. Hämäläinen, "Assessing and improving the spatial accuracy in meg source localization by depth-weighted minimum-norm estimates," *NeuroImage*, vol. 31, no. 1, pp. 160–171, 2006.
- [13] W. J. Freeman, M. D. Holmes, B. C. Burke, and S. Vanhatalo, "Spatial spectra of scalp EEG and EMG from awake humans," *Clinical Neurophysiology*, vol. 114, no. 6, pp. 1053–1068, 2003.
- [14] M. Odabae, W. J. Freeman, P. B. Colditz, C. Ramon, and S. Vanhatalo, "Spatial patterning of the neonatal EEG suggests a need for a high number of electrodes," *NeuroImage*, vol. 68, pp. 229–235, 2013.
- [15] C. Ramon, M. Holmes, W. J. Freeman, M. Gratkowski, K. Eriksen, and J. Haueisen, "Power spectral density changes and language lateralization during covert object naming tasks measured with high-density EEG recordings," *Epilepsy & Behavior*, vol. 14, no. 1, pp. 54–59, 2009.
- [16] Y. Petrov, J. Nador, C. Hughes, S. Tran, O. Yavuzcetin, and S. Sridhar, "Ultra-dense EEG sampling results in two-fold increase of functional brain information," *NeuroImage*, vol. 90, pp. 140–145, 2014.
- [17] Texas Instruments, "ADS 1299: Low-Noise, 8-Channel, 24-Bit Analog Front-End for Biopotential Measurements," 2015. [Online]. Available: <http://www.ti.com/product/ads1299>
- [18] K. T. Kim and T. Berger, "Sending a lossy version of the innovations process is suboptimal in QG rate-distortion," in *IEEE International Symposium on Information Theory (ISIT)*, 2005, pp. 209–213.
- [19] —, "The degree of suboptimality of sending a lossy version of the innovations process in Gauss-Markov rate-distortion," in *IEEE International Symposium on Information Theory (ISIT)*, 2006, pp. 808–812.
- [20] R. Zamir, Y. Kochman, and U. Erez, "Achieving the Gaussian rate–distortion function by prediction," *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 3354–3364, 2008.
- [21] K. F. Riley and M. P. Hobson, *Essential mathematical methods for the physical sciences*. Cambridge University Press, 2011.
- [22] Y. Zhang, W. van Drongelen, and B. He, "Estimation of in vivo brain-to-skull conductivity ratio in humans," *Applied physics letters*, vol. 89, no. 22, 2006.
- [23] R. Srinivasan, P. L. Nunez, D. M. Tucker, R. B. Silberstein, and P. J. Cadusch, "Spatial sampling and filtering of EEG with spline laplacians to estimate cortical potentials," *Brain topography*, vol. 8, no. 4, pp. 355–366, 1996.
- [24] "Code for a case study on power savings for high-density EEG using a hierarchical measurement technique." [Online]. Available: <http://www.tinyurl.com/pulkitgrover/EEGCode.html>
- [25] Z. Khalid, R. A. Kennedy, and J. D. McEwen, "An optimal-dimensionality sampling scheme on the sphere with fast spherical harmonic transforms," *Signal Processing, IEEE Transactions on*, vol. 62, no. 17, pp. 4597–4610, 2014.
- [26] J. D. McEwen and Y. Wiaux, "A novel sampling theorem on the sphere," *Signal Processing, IEEE Transactions on*, vol. 59, no. 12, pp. 5876–5887, 2011.
- [27] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 67, no. 2, pp. 301–320, 2005.
- [28] R. D. Pascual-Marqui, "Review of methods for solving the EEG inverse problem," *International journal of bioelectromagnetism*, vol. 1, no. 1, pp. 75–86, 1999.
- [29] M. Tegmark, "An icosahedron-based method for pixelizing the celestial sphere," *The Astrophysical Journal*, vol. 470, p. L81, 1996.
- [30] A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonen *et al.*, "MEG and EEG data analysis with MNE-Python," 2013.
- [31] J. Bae, A. Deshmukh, Y. Song, and J. Riera, "Brain Source Imaging in Preclinical Rat Models of Focal Epilepsy using High-Resolution EEG Recordings," *Journal of visualized experiments: JoVE*, no. 100, 2015.
- [32] C. Guger, G. Krausz, B. Z. Allison, and G. Edlinger, "Comparison of dry and gel based electrodes for p300 brain-computer interfaces," *Front Neurosci*, vol. 6, no. 60, pp. 1–7, 2012.
- [33] P. Grover, "Fundamental limits on source-localization accuracy of EEG-based brain interfaces," in *IEEE International Symposium on Information Theory (ISIT)*, Barcelona, Spain, 2016.
- [34] S. Baillet, J. C. Mosher, and R. M. Leahy, "Electromagnetic brain mapping," *IEEE Signal Processing Magazine*, vol. 18, no. 6, pp. 14–30, 2001.
- [35] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York: Wiley, 2006.
- [36] G. Lantz, R. G. de Peralta, L. Spinelli, M. Seeck, and C. Michel, "Epileptic source localization with high density eeg: how many electrodes are needed?" *Clinical neurophysiology*, vol. 114, no. 1, pp. 63–69, 2003.
- [37] P. Grover, J. Weldon, S. Kelly, P. Venkatesh, and H. Jeong, "A Novel Information-Theoretic Sensing Strategy for Ultra High-Density EEG Sensing," in *Proceedings of the Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Oct. 2015.
- [38] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, "A review of classification algorithms for eeg-based brain–computer interfaces," *Journal of neural engineering*, vol. 4, no. 2, p. R1, 2007.
- [39] M. A. Lebedev and M. A. Nicolelis, "Brain-machine interfaces: past, present and future," *TRENDS in Neurosciences*, vol. 29, no. 9, pp. 536–546, 2006.
- [40] T. Han and S.-I. Amari, "Statistical inference under multiterminal data compression," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2300–2324, Oct 1998.
- [41] T. Berger, "Decentralized estimation and decision theory," *presented at the IEEE 7th Workshop on Information Theory*, NY, Sept 1979.
- [42] M. Mahzoon, H. Albalawi, X. Li, and P. Grover, "Using relative-relevance of data pieces for efficient communication, with an application to neural data acquisition," in *52nd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2014, pp. 160–166.
- [43] M. Mahzoon, C. Li, X. Li, and P. Grover, "Energy-constrained distributed learning and classification by exploiting relative relevance of sensors' data," *IEEE J. Sel. Areas Commun.*, no. 5, pp. 1417–1430, May 2016.
- [44] S. Du, Y. Xu, H. Zhang, C. Li, P. Grover, and A. Singh, "Novel quantization strategies for linear prediction with guarantees," in *International Conference on Machine Learning (ICML), Workshop on "on-device intelligence"*, New York, USA, 2016.
- [45] A. D. Wyner and J. Ziv, "The Rate-Distortion Function for Source Coding with Side Information at the Decoder," *IEEE Trans. Inf. Theory*, vol. 22, no. 1, p. 1, 1976.
- [46] A. Cayley, "A theorem on trees," *Quarterly Journal of Pure and Applied Mathematics*, vol. 23, pp. 376–378, 1889.
- [47] M. J. Boring, S. Kelly, J. Weldon, A. Robinson, M. Behrmann, and P. Grover, "Containing errors in computations for neural sensing: does a hierarchical-referencing strategy lead to energy savings?" *Workshop on Information Theory and its Applications (ITA)*, 2016.
- [48] P. Grover, "Ultra-high-density EEG: how many bits of resolution do the electrodes need?" in *Asilomar conference on Signals, systems and computers*, Pacific Grove, CA, USA, 2015.
- [49] B. N. Cuffin, D. L. Schomer, J. R. Ives, and H. Blume, "Experimental tests of EEG source localization accuracy in spherical head models," *Clinical Neurophysiology*, vol. 112, no. 1, pp. 46–51, 2001.
- [50] R. Hoekema, G. Wieneke, F. Leijten, C. Van Veen, P. Van Rijen, G. Huiskamp, J. Ansems, and A. Van Huffelen, "Measurement of the conductivity of skull, temporarily removed during epilepsy surgery," *Brain topography*, vol. 16, no. 1, pp. 29–38, 2003.
- [51] A. El Gamal and Y.-H. Kim, *Network information theory*. Cambridge university press, 2011.
- [52] S. Kim, C. J. Quinn, N. Kiyavash, and T. P. Coleman, "Dynamic and succinct statistical analysis of neuroscience data," *Proceedings of the IEEE*, vol. 102, no. 5, pp. 683–698, 2014.
- [53] P. Venkatesh and P. Grover, "Is the direction of greater Granger causal influence same as the direction of information flow?" in *Proceedings of the Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Oct. 2015.

APPENDIX A TRANSFER FUNCTION COMPUTATION

The boundary conditions are:

$$\Phi^{(s-1)} \Big|_{r=r_{s-1}} = \Phi^{(s)} \Big|_{r=r_{s-1}} . \quad (17)$$

$$\sigma_{s-1} \frac{\partial \Phi^{(s-1)}}{\partial r} \Big|_{r=r_{s-1}} = \sigma_s \frac{\partial \Phi^{(s)}}{\partial r} \Big|_{r=r_{s-1}} , \quad (18)$$

for the N -th (last) sphere, the current dies down outside it (because the conductivity of air, σ_{N+1} is 0), yielding:

$$\sigma_N \frac{\partial \Phi^{(N)}}{\partial r} \Big|_{r=r_N} = 0 \quad (19)$$

Now, from (17), we get:

$$\begin{aligned} \Phi^{(s-1)} \Big|_{r=r_{s-1}} = \Phi^{(s)} \Big|_{r=r_{s-1}} &\Rightarrow A_{lm}^{(s-1)} \left(\frac{r}{r_{s-1}} \right)^l + B_{lm}^{(s-1)} \left(\frac{r_{s-1}}{r} \right)^{l+1} \Big|_{r=r_{s-1}} = A_{lm}^{(s)} \left(\frac{r}{r_s} \right)^l + B_{lm}^{(s)} \left(\frac{r_s}{r} \right)^{l+1} \Big|_{r=r_{s-1}} \\ &\Rightarrow A_{lm}^{(s-1)} + B_{lm}^{(s-1)} = A_{lm}^{(s)} \left(\frac{r_{s-1}}{r_s} \right)^l + B_{lm}^{(s)} \left(\frac{r_s}{r_{s-1}} \right)^{l+1} \end{aligned} \quad (20)$$

Finally, from (18), we get:

$$\begin{aligned} \sigma_{s-1} \frac{\partial \Phi^{(s-1)}}{\partial r} \Big|_{r=r_{s-1}} &= \sigma_s \frac{\partial \Phi^{(s)}}{\partial r} \Big|_{r=r_{s-1}} \\ &\Rightarrow \sigma_{s-1} \left(A_{lm}^{(s-1)} \frac{lr^{l-1}}{r_{s-1}^l} - B_{lm}^{(s-1)} (l+1) \frac{r_{s-1}^{l+1}}{r^{l+2}} \right) \Big|_{r=r_{s-1}} = \sigma_s \left(A_{lm}^{(s)} \frac{lr^{l-1}}{r_s^l} - B_{lm}^{(s)} (l+1) \frac{r_s^{l+1}}{r^{l+2}} \right) \Big|_{r=r_{s-1}} \\ &\Rightarrow \sigma_{s-1} \left(A_{lm}^{(s-1)} \frac{lr_{s-1}^{l-1}}{r_{s-1}^l} - B_{lm}^{(s-1)} (l+1) \frac{r_{s-1}^{l+1}}{r_{s-1}^{l+2}} \right) = \sigma_s \left(A_{lm}^{(s)} \frac{lr_{s-1}^{l-1}}{r_s^l} - B_{lm}^{(s)} (l+1) \frac{r_s^{l+1}}{r_{s-1}^{l+2}} \right) \\ &\Rightarrow \sigma_{s-1} \left(A_{lm}^{(s-1)} \frac{l}{r_{s-1}} - B_{lm}^{(s-1)} (l+1) \frac{1}{r_{s-1}} \right) = \sigma_s \left(A_{lm}^{(s)} \frac{lr_{s-1}^{l-1}}{r_s^l} - B_{lm}^{(s)} (l+1) \frac{r_s^{l+1}}{r_{s-1}^{l+2}} \right) \\ &\Rightarrow lA_{lm}^{(s-1)} - (l+1)B_{lm}^{(s-1)} = \sigma_{s,s-1} \left(lA_{lm}^{(s)} \left(\frac{r_{s-1}}{r_s} \right)^l - (l+1)B_{lm}^{(s)} \left(\frac{r_s}{r_{s-1}} \right)^{l+1} \right) \end{aligned} \quad (21)$$

From (20), assuming $B_{lm}^{(s)} = \gamma_l^{(s)} A_{lm}^{(s)}$,

$$\begin{aligned} A_{lm}^{(s-1)} + B_{lm}^{(s-1)} &= A_{lm}^{(s)} \left(\frac{r_{s-1}}{r_s} \right)^l + B_{lm}^{(s)} \left(\frac{r_s}{r_{s-1}} \right)^{l+1} = A_{lm}^{(s)} \left(\left(\frac{r_{s-1}}{r_s} \right)^l + \gamma_l^{(s)} \left(\frac{r_s}{r_{s-1}} \right)^{l+1} \right) \\ &\Rightarrow A_{lm}^{(s)} = \frac{A_{lm}^{(s-1)} + B_{lm}^{(s-1)}}{\left(\left(\frac{r_{s-1}}{r_s} \right)^l + \gamma_l^{(s)} \left(\frac{r_s}{r_{s-1}} \right)^{l+1} \right)} \end{aligned} \quad (22)$$

Now, from (21), again assuming $B_{lm}^{(s)} = \gamma_l^{(s)} A_{lm}^{(s)}$,

$$\begin{aligned} lA_{lm}^{(s-1)} - (l+1)B_{lm}^{(s-1)} &= \sigma_{s,s-1} \left(lA_{lm}^{(s)} \left(\frac{r_{s-1}}{r_s} \right)^l - (l+1)B_{lm}^{(s)} \left(\frac{r_s}{r_{s-1}} \right)^{l+1} \right) \\ &= \sigma_{s,s-1} \left(l \left(\frac{r_{s-1}}{r_s} \right)^l - (l+1)\gamma_l^{(s)} \left(\frac{r_s}{r_{s-1}} \right)^{l+1} \right) A_{lm}^{(s)} \\ &= \sigma_{s,s-1} \left(l \left(\frac{r_{s-1}}{r_s} \right)^l - (l+1)\gamma_l^{(s)} \left(\frac{r_s}{r_{s-1}} \right)^{l+1} \right) \times \frac{A_{lm}^{(s-1)} + B_{lm}^{(s-1)}}{\left(\left(\frac{r_{s-1}}{r_s} \right)^l + \gamma_l^{(s)} \left(\frac{r_s}{r_{s-1}} \right)^{l+1} \right)} = \sigma_{s,s-1} \zeta_{ls} \left(A_{lm}^{(s-1)} + B_{lm}^{(s-1)} \right) \\ &\Rightarrow A_{lm}^{(s-1)} (l - \sigma_{s,s-1} \zeta_{ls}) = B_{lm}^{(s-1)} ((l+1) + \sigma_{s,s-1} \zeta_{ls}) \Rightarrow B_{lm}^{(s-1)} = \gamma_l^{(s-1)} A_{lm}^{(s-1)}, \end{aligned}$$

where

$$\gamma_l^{(s-1)} = \frac{l - \sigma_{s,s-1} \zeta_{ls}}{(l+1) + \sigma_{s,s-1} \zeta_{ls}} = \frac{l \sigma_{s-1,s} - \zeta_{ls}}{(l+1) \sigma_{s-1,s} + \zeta_{ls}} = \frac{\frac{l}{l+1} \sigma_{s-1,s} - \frac{\zeta_{ls}}{l+1}}{\sigma_{s-1,s} + \frac{\zeta_{ls}}{l+1}}, \text{ and } \zeta_{ls} = \frac{l \left(\frac{r_{s-1}}{r_s} \right)^l - (l+1) \gamma_l^{(s)} \left(\frac{r_s}{r_{s-1}} \right)^{l+1}}{\left(\left(\frac{r_{s-1}}{r_s} \right)^l + \gamma_l^{(s)} \left(\frac{r_s}{r_{s-1}} \right)^{l+1} \right)}$$

To start the induction, from (19), we get:

$$\begin{aligned} \sigma_N \frac{\partial \Phi^{(N)}}{\partial r} \Big|_{r=r_N} &= 0 \Rightarrow \sum_{l=1}^{\infty} \left[A_{lm}^{(N)} \frac{lr^{l-1}}{r_N^l} - B_{lm}^{(N)} (l+1) \frac{r_N^{l+1}}{r_N^{l+2}} \right] Y_{lm}(\theta, \phi) \Big|_{r=r_N} = 0, \\ \Rightarrow \sum_{l=1}^{\infty} \left[A_{lm}^{(N)} \frac{l}{r_N} - B_{lm}^{(N)} (l+1) \frac{1}{r_N} \right] Y_{lm}(\theta, \phi) \Big|_{r=r_N} &= 0 \Rightarrow B_{lm}^{(N)} = \frac{l}{l+1} A_{lm}^{(N)} = \gamma_l^{(N)} A_{lm}^{(N)}, \end{aligned}$$

where $\gamma_l^{(N)} = \frac{l}{l+1}$. Thus the induction assumption is satisfied, and we can use (22) and (23) to continue the induction (i.e., go from $s = N$ to $s = N-1, N-2, \dots, 1$).

APPENDIX B APPROXIMATION ANALYSIS

We know that $B_{lm}^{(s)} = A_{lm}^{(s)} \gamma_l^{(s)}$, and thus $c_{l,m}^{(s)} = A_{lm}^{(s)} + B_{lm}^{(s)} = (1 + \gamma_l^{(s)}) A_{lm}^{(s)}$. Thus,

$$\begin{aligned} c_{l,m}^{(s)} &= (1 + \gamma_l^{(s)}) A_{lm}^{(s)} = (1 + \gamma_l^{(s)}) \frac{A_{lm}^{(s-1)} + B_{lm}^{(s-1)}}{\left(\left(\frac{r_{s-1}}{r_s} \right)^l + \gamma_l^{(s)} \left(\frac{r_s}{r_{s-1}} \right)^{l+1} \right)} = (1 + \gamma_l^{(s)}) \frac{c_{l,m}^{(s-1)}}{\left(\left(\frac{r_{s-1}}{r_s} \right)^l + \gamma_l^{(s)} \left(\frac{r_s}{r_{s-1}} \right)^{l+1} \right)} \\ &\stackrel{(l \text{ large})}{\approx} \frac{(1 + \gamma_l^{(s)})}{\gamma_l^{(s)}} \frac{c_{l,m}^{(s-1)}}{\left(\frac{r_s}{r_{s-1}} \right)^{l+1}} \end{aligned} \quad (23)$$

Thus, solving for $c_{l,m}^{(N)}$,

$$c_{l,m}^{(N)} \approx \left(\prod_{s=2}^N \frac{(1 + \gamma_l^{(s)})}{\gamma_l^{(s)}} \right) \frac{1}{\left(\frac{r_N}{r_{N-1}} \frac{r_{N-1}}{r_{N-2}} \dots \frac{r_2}{r_1} \right)^{l+1}} c_{l,m}^{(1)} = \left(\prod_{s=2}^N \frac{(1 + \gamma_l^{(s)})}{\gamma_l^{(s)}} \right) \frac{1}{\left(\frac{r_N}{r_1} \right)^{l+1}} c_{l,m}^{(1)}. \quad (24)$$

The limits of $\gamma_l^{(s)}$ as $l \rightarrow \infty$ can be obtained as follows: We will evaluate the constants γ and ξ for a fixed s as $l \rightarrow \infty$. As $l \rightarrow \infty$, $\gamma_l^N \rightarrow 1$. For $s = 2, \dots, N$, it is easy to see that $\frac{\xi_{l,s}}{l+1} \rightarrow -1$. Thus, for $s = 2, \dots, N-1$, $\gamma_l^{(s)} \rightarrow \frac{\sigma_{s,s+1}+1}{\sigma_{s,s+1}-1}$. The result now follows from substituting the limits of $\gamma_l^{(s)}$ and observing for the N -th term that σ_{N+1} is 0, the conductivity of air.