
This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.

Iivanainen, Joonas; Mäkinen, Antti J.; Zetter, Rasmus; Stenroos, Matti; Ilmoniemi, Risto J.; Parkkonen, Lauri

Spatial sampling of MEG and EEG based on generalized spatial-frequency analysis and optimal design

Published in:
NeuroImage

DOI:
[10.1016/j.neuroimage.2021.118747](https://doi.org/10.1016/j.neuroimage.2021.118747)

Published: 15/12/2021

Document Version
Publisher's PDF, also known as Version of record

Published under the following license:
CC BY-NC-ND

Please cite the original version:
Iivanainen, J., Mäkinen, A. J., Zetter, R., Stenroos, M., Ilmoniemi, R. J., & Parkkonen, L. (2021). Spatial sampling of MEG and EEG based on generalized spatial-frequency analysis and optimal design. *NeuroImage*, 245, 1-15. Article 118747. <https://doi.org/10.1016/j.neuroimage.2021.118747>



Spatial sampling of MEG and EEG based on generalized spatial-frequency analysis and optimal design

Joonas Iivanainen^{1,*}, Antti J. Mäkinen¹, Rasmus Zetter, Matti Stenroos, Risto J. Ilmoniemi, Lauri Parkkonen

Department of Neuroscience and Biomedical Engineering, Aalto University School of Science, Aalto FI-00076, Finland

ARTICLE INFO

Keywords:

Magnetoencephalography
Electroencephalography
On-scalp MEG
Spatial sampling
Optimal design
Spatial frequency

ABSTRACT

In this paper, we analyze spatial sampling of electro- (EEG) and magnetoencephalography (MEG), where the electric or magnetic field is typically sampled on a curved surface such as the scalp. By simulating fields originating from a representative adult-male head, we study the spatial-frequency content in EEG as well as in on- and off-scalp MEG. This analysis suggests that on-scalp MEG, off-scalp MEG and EEG can benefit from up to 280, 90 and 110 spatial samples, respectively. In addition, we suggest a new approach to obtain sensor locations that are optimal with respect to prior assumptions. The approach also allows to control, e.g., the uniformity of the sensor locations. Based on our simulations, we argue that for a low number of spatial samples, model-informed non-uniform sampling can be beneficial. For a large number of samples, uniform sampling grids yield nearly the same total information as the model-informed grids.

1. Introduction

Electro- (EEG) and magnetoencephalography (MEG) are noninvasive neuroimaging techniques for measuring brain activity at millisecond time scale (Hämäläinen et al., 1993; Nunez et al., 2006). EEG and MEG measure the electric and magnetic field, respectively, due to neuronal currents. Adequate spatial sampling is necessary to capture the spatial detail of the continuous field due to brain activity. EEG setups typically involve 16–128 electrodes placed uniformly on the subject's scalp while state-of-the-art MEG systems use around 300 superconducting quantum interference device (SQUID) sensors placed rigidly around the subject's head (Baillet, 2017).

Spatial sampling of EEG and MEG has again become a topic of interest. High-density EEG (hdEEG) with an electrode count of 128–256 or even more has been suggested to be beneficial (e.g., Brodbeck et al. 2011; Grover and Venkatesh 2016; Hedrich et al. 2017; Petrov et al. 2014; Robinson et al. 2017). In MEG, in contrast to conventional liquid-helium-cooled SQUID sensors that measure the field ~2 cm off the scalp, novel sensors such as optically-pumped magnetometers (OPMs; Budker and Romalis 2007) and high- T_c SQUIDs (Faley et al., 2017) have enabled on-scalp field sensing within millimetres from the head surface. Simulations have shown that the closer proximity of the sensors to the brain improves spatial resolution of MEG and provides more information about neuronal currents (Boto et al., 2016; Iivanainen et al., 2017;

Riaz et al., 2017; Schneiderman, 2014). The number of sensors that is beneficial in on-scalp MEG is currently not known well. A recent simulation study suggested that fewer sensors are needed in on-scalp MEG than with SQUIDS to achieve similar spatial discrimination performance (Tierney et al., 2019), while another experimental study suggested that ~50 OPMs give comparable results to a state-of-the-art SQUID system (Hill et al., 2020).

The number of sensors or, equivalently, sensor spacing in EEG has been extensively studied. Srinivasan and colleagues (1998) argued that to characterize the full range of spatial detail available for EEG, at least 128 electrodes are needed (spacing ~3 cm). More recently, finite-element modeling in a realistic head model suggested a minimum electrode spacing of 50–59 mm determined by the distance at which the potential decreased to 10% of its peak (Slutzky et al., 2010). Experiments by Freeman and colleagues (2003) suggested that electrode spacing as small as 5–8 mm could be beneficial, if high spatial frequencies contain behaviorally relevant information above the noise level.

The sensor spacing in MEG has been less studied. Analytical calculations in half-infinite homogeneous volume conductors suggest that the sensor spacing should be approximately equal to the distance of the sensors to the brain (Ahonen et al., 1993). In many MEG studies, different sensor arrays have been compared without necessarily formulating the comparison as a sampling problem (Boto et al., 2016; Iivanainen et al., 2017; Nenonen et al., 2004; Riaz et al., 2017; Wilson and Vrba, 2007).

* Corresponding author.

E-mail addresses: jaiivanainen@gmail.com (J. Iivanainen), makinen.antti.j@gmail.com (A.J. Mäkinen).

¹ These authors contributed equally to this work.

Previous studies have mainly focused on spatial sampling where the sensors cover the entire scalp uniformly. However, the field can also be sampled nonuniformly. In certain applications, such as brain-computer interfaces, targeted non-uniform sampling would be of value: one could reduce the sensor count while maintaining spatial resolution and sensitivity to the cortical area of interest.

In this work, we analyze spatial sampling of EEG and MEG on realistically curved surfaces. We carry out a spatial-frequency analysis of continuous fields using the eigenfunctions of the surface Laplace–Beltrami operator (Bronstein et al., 2017; Reuter et al., 2009). The basis formed by these functions can be seen as a natural generalization of the 1-D Fourier basis on a surface. These functions have been used in a variety of applications from signal and geometry processing (Levy, 2006; Reuter et al., 2009) to cortical analysis (Qiu et al., 2006) and deep learning (Bronstein et al., 2017). We further describe how more compact bases can be constructed using prior information of the neuronal field patterns. We discuss how these basis-function representations of the field patterns relate to the number of spatial samples.

We investigate how to obtain optimal sample positions in EEG and MEG. We utilize Gaussian processes (Abrahamsen, 1997; Williams and Rasmussen, 2006) to encode prior knowledge about the field. This perspective is similar to kriging in geospatial sciences (Chilès and Desassis, 2018; Cressie, 1993) and Gaussian-process regression in machine learning (Williams and Rasmussen, 2006). We introduce measures from experimental design (Chaloner and Verdinelli, 1995; Krause et al., 2008; Lindley, 1956) that can be used to quantify the optimality of the sample positions. We suggest a method that maximizes Shannon’s information (Lindley, 1956; Shannon, 1949) for a given number of samples and prior assumptions. We use the method to generate sampling grids for different EEG and MEG experiments, where either the whole brain or parts of it are of interest.

2. Theory and methods

In this Section, we introduce the theoretical concepts behind our analysis of sampling of neuro-electromagnetic fields. In Section 2.1, we describe how a scalar field measured on a surface can be analyzed using a spatial-frequency basis. In Section 2.2, we introduce the spatial-frequency-limited representation of the field and analyze the reconstruction error due to the use of such a representation. This analysis provides the motivation for using the 99%-energy threshold for the number of spatial-frequency components in the field, which we later use to quantify the number of samples beneficial in spatial sampling of MEG/EEG.

In Section 2.3, we change the perspective to random fields, which provide another framework for field reconstruction and insight to the metrics we use to quantify the performance of a sampling grid in Section 2.5. The concept of a field kernel is introduced in Section 2.3 and further described in Section 2.4; we employ such a kernel in constructing optimized sampling grids in Section 2.6.

2.1. Spatial-frequency representation of dipole fields

The quasi-static electric potential or magnetic field component $f(\vec{r})$ is generated by distributed source activity $\vec{q}(\vec{r}_s)$ inside the brain and can be written in terms of Green’s function $\vec{L}(\vec{r}, \vec{r}_s)$ as (Hämäläinen et al., 1993; Nunez et al., 2006)

$$f(\vec{r}) = \int \vec{L}(\vec{r}, \vec{r}_s) \cdot \vec{q}(\vec{r}_s) dV_s. \quad (1)$$

By discretizing the neural source distribution $\vec{q}(\vec{r}_s)$ into a set of primary current dipoles $\vec{q}_i(\vec{r}_i) = q_i \hat{q}_i \delta(\vec{r} - \vec{r}_i)$, $i = 1, \dots, M$, where q_i is the amplitude of the i th source at \vec{r}_i and \hat{q}_i its orientation, Eq. (1) becomes

$$f(\vec{r}) = \sum_i \left(\vec{L}(\vec{r}, \vec{r}_i) \cdot \hat{q}_i \right) q_i = \sum_i f_i(\vec{r}) q_i, \quad (2)$$

i.e., the field is composed of *dipole fields* $f_i(\vec{r}) = \vec{L}(\vec{r}, \vec{r}_i) \cdot \hat{q}_i$ corresponding to unit dipole sources. When the field is also discretized to (or sampled

at) N points, the equation reduces to a linear matrix equation $\mathbf{f} = \mathbf{L}\mathbf{q}$, where \mathbf{L} is the $(N \times M)$ lead-field matrix (Hämäläinen et al., 1993). Please note that with an arrow we denote a Euclidean vector in the 3-D space while bolded letters refer to column vectors and matrices.

In MEG and EEG, we sample $f(\vec{r})$ on a curved 2-D surface embedded in 3-D space. Conventional 1-D Fourier analysis can be extended to spatial-frequency (SF) analysis on such surfaces using a suitable orthonormal function basis. The generating equation for the spatial-frequency basis $\{u_m\}$ is the Helmholtz equation (Levy, 2006)

$$-\nabla_{\text{LB}}^2 u_m = k_m^2 u_m, \quad (3)$$

which is the eigenvalue equation of the Laplace–Beltrami (LB) operator ∇_{LB}^2 , i.e., the surface Laplacian. The eigenfunctions u_m are modes of standing waves on the surface, with increasing spatial frequency and complexity towards higher m (Fig. 3B). The square roots of the eigenvalues k_m generalize the concept of wavenumber. For example, in 1-D, k_m correspond to the angular wavenumber of a sinusoid. Generalizing this relationship for eigenfunctions on a surface, the wavenumbers k_m correspond to wavelengths approximately as $2\pi/k_m$. On a compact surface, the eigenvalue spectrum is discrete $\{k_m, u_m\}$ and the eigenfunctions form an orthonormal basis with respect to the inner product (Levy, 2006):

$$\langle u_m, u_{m'} \rangle = \int_S u_m(\vec{r}) u_{m'}(\vec{r}) dS = \delta_{m,m'}, \quad (4)$$

where $\delta_{m,m'}$ is the Kronecker delta.

The SF basis can be used to analyze the spatial-frequency content of the fields. As the basis is orthonormal, any $f(\vec{r})$ can be expressed as a linear combination of SF basis functions

$$f(\vec{r}) = \sum_{m=1}^{\infty} a_m u_m(\vec{r}) = \sum_{m=1}^{\infty} \langle f, u_m \rangle u_m(\vec{r}), \quad (5)$$

where the coefficients are projections of $f(\vec{r})$ on $\{u_m\}$. Due to orthonormality, the field energy can be decomposed as

$$\|f\|^2 = \langle f, f \rangle = \int_S |f(\vec{r})|^2 dS = \sum_{m=1}^{\infty} a_m^2, \quad (6)$$

a relation similar to Parseval’s theorem for Fourier series. The squared coefficients $a_m^2 = \langle f, u_m \rangle^2$ comprise the *energy spectrum* of the field in the sense of spectral signal analysis.

Sampling theorems typically assume that the signals are bandlimited [e.g., in 1-D (Shannon, 1949); on a sphere (McEwen and Wiaux, 2011); on a surface/manifold (Pesenson, 2014; 2015)]. In our context, a bandlimited field can be expressed with a limited number of SF basis functions $\{u_m\}$, i.e., $a_m = 0$ for $m > B$ ($k_m > k_B$). On a surface, the sampling theorem by Pesenson (2015) states that a k_B -bandlimited field is uniquely determined by its values on a grid with a uniform sample spacing $\rho = ck_B^{-1}$, where c is a proportionality constant that depends on the surface geometry.

Figure 1 shows a spatial-frequency representation of a dipole field in MEG and EEG. Although most of the field energy in general resides at low spatial frequencies, fields are not strictly bandlimited. Thereby, standard sampling theorems are not directly applicable.

2.2. Sampling and reconstruction assuming a bandlimit

Here, we analyze how an effective bandlimit B for the field can be chosen by studying reconstruction of f from noisy samples. This section introduces the methodology used in Results Section 3.2.

The field can be represented as

$$f(\vec{r}) = f_B(\vec{r}) + f_r(\vec{r}) = \sum_{m=1}^B a_m u_m(\vec{r}) + \sum_{m=B+1}^{\infty} a_m u_m(\vec{r}), \quad (7)$$

where f_B is the B -bandlimited representation of the field and f_r the residual. We assume that the field can be approximated with the B first SF basis functions as $f \approx f_B$. To simplify the notation of Eq. (7), we

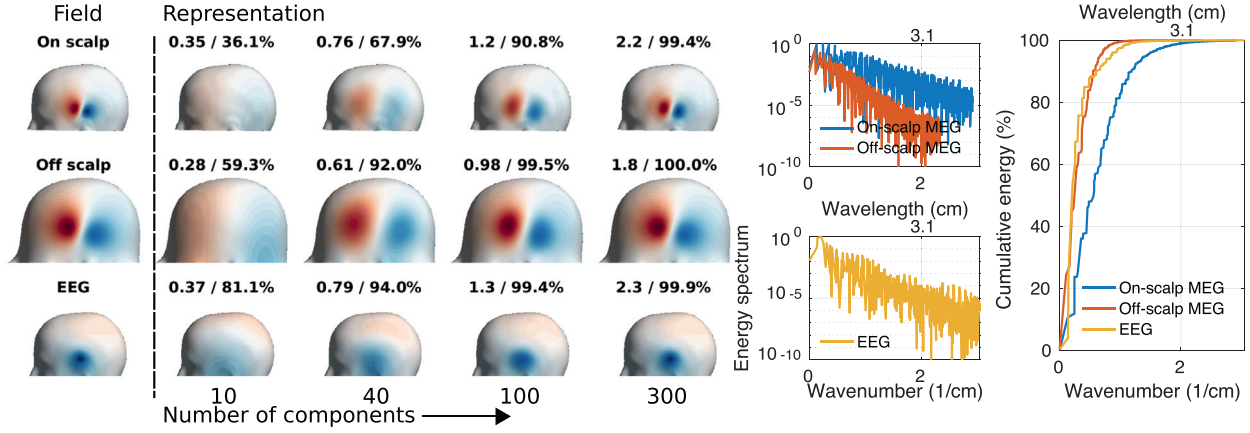


Fig. 1. Spatial-frequency representations of dipole fields in MEG and EEG. ‘On scalp’ refers to MEG measurement within millimetres from the scalp while ‘off scalp’ to that within 2 cm. For details, see Section 3.1. **Left:** Representation of a dipole field with an increasing number of spatial-frequency components. The maximum wavenumber (1/cm) and the cumulative energy of the representation are shown. **Right:** Energy spectrum and cumulative energy of the field as a function of wavenumber and wavelength (upper x-axis; estimated from the wavenumber k as $2\pi/k$).

express the summation over the basis functions as a dot product $f_B(\vec{r}) = \mathbf{u}(\vec{r})^T \mathbf{a}$ where $\mathbf{u}(\vec{r})[m] = u_m(\vec{r})$ and $\mathbf{a}[m] = a_m$.

We model N noisy samples of $f_B(\vec{r})$ at locations \vec{r}_n stacked in a column vector \mathbf{y} as

$$\mathbf{y} = \begin{bmatrix} \mathbf{u}(\vec{r}_1)^T \\ \vdots \\ \mathbf{u}(\vec{r}_N)^T \end{bmatrix} \mathbf{a} + \boldsymbol{\epsilon} = \mathbf{U}\mathbf{a} + \boldsymbol{\epsilon}, \quad (8)$$

where \mathbf{U} is an $(N \times B)$ matrix with the column vectors \mathbf{u} at \vec{r}_n on rows, i.e., $\mathbf{U}[n, m] = u_m(\vec{r}_n)$, and $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \sigma^2 \mathbf{I})$ is white measurement noise. When $N \geq B$, we can estimate the coefficients \mathbf{a} optimally using the least-squares method (Kailath et al., 2000)

$$\hat{\mathbf{a}} = (\mathbf{U}^T \mathbf{U})^{-1} \mathbf{U}^T \mathbf{y} = \mathbf{G}\mathbf{y}, \quad (9)$$

and the bandlimited field f_B can be reconstructed as

$$\hat{f}_B(\vec{r}) = \mathbf{u}(\vec{r})^T \hat{\mathbf{a}} = \mathbf{u}(\vec{r})^T \mathbf{G}\mathbf{y} = \sum_n \left(\sum_m \mathbf{G}[m, n] u_m(\vec{r}) \right) y_n, \quad (10)$$

where the functions $\sum_m \mathbf{G}[m, n] u_m(\vec{r})$ are interpolation functions for the samples y_n .

To study the reconstruction error, we update the measurement model as $\mathbf{y} = \mathbf{U}\mathbf{a} + \mathbf{U}_r \mathbf{a}_r + \boldsymbol{\epsilon}$, where $\mathbf{U}_r[n, m] = u_{m+B}(\vec{r}_n)$ corresponds to the residual field f_r . Inserting \mathbf{y} into Eq. (9), the coefficient estimate can be expressed as

$$\hat{\mathbf{a}} = \mathbf{G}(\mathbf{U}\mathbf{a} + \mathbf{U}_r \mathbf{a}_r + \boldsymbol{\epsilon}) = \mathbf{a} + \mathbf{G}(\mathbf{U}_r \mathbf{a}_r + \boldsymbol{\epsilon}) = \mathbf{a} + \Delta \mathbf{a}, \quad (11)$$

where $\Delta \mathbf{a} = \mathbf{G}(\mathbf{U}_r \mathbf{a}_r + \boldsymbol{\epsilon})$ is the error in the estimated coefficients. The error consists of two parts: $\mathbf{G}\mathbf{U}_r \mathbf{a}_r$ is the error due to aliasing from components outside the band $m \leq B$, and $\mathbf{G}\boldsymbol{\epsilon}$ is the random error due to noise.

The expected reconstruction error can now be evaluated as

$$\mathbb{E} \|\mathbf{f} - \hat{\mathbf{f}}_B\|^2 = \mathbb{E} \|\mathbf{f}_B - \hat{\mathbf{f}}_B\|^2 + \mathbb{E} \|\mathbf{f}_r\|^2 = \mathbb{E} \|\Delta \mathbf{a}\|^2 + \mathbb{E} \|\mathbf{a}_r\|^2. \quad (12)$$

Since $\mathbb{E}(\|\mathbf{G}\boldsymbol{\epsilon}\|^2) = \text{Tr}[\sigma^2(\mathbf{U}^T \mathbf{U})^{-1}]$ and $\mathbb{E}(\boldsymbol{\epsilon}) = 0$, we can rewrite the error as

$$\mathbb{E} \|\mathbf{f} - \hat{\mathbf{f}}\|^2 = \text{Tr}[\sigma^2(\mathbf{U}^T \mathbf{U})^{-1}] + \mathbb{E}(\|\mathbf{G}\mathbf{U}_r \mathbf{a}_r\|^2) + \mathbb{E}(\|\mathbf{a}_r\|^2). \quad (13)$$

If the noise term dominates the reconstruction error, f_B can be considered a reasonably good approximation of f ; the white noise variance σ^2 can then be used to determine the number of components B and the number of samples $N \geq B$ needed for the reconstruction. If the noise level is unknown, a threshold such as 1% expected residual can be used to determine the effective bandlimit B . This threshold may be obtained by finding the B that covers at least 99% energy of every dipole

field f_i . This bounds the expected error of a random distribution of such sources (Appendix A). A similar threshold has been used previously (e.g., Ahonen et al. 1993; Grover and Venkatesh 2016). Last, both the noise and the aliasing error depend on the sample positions $\{\vec{r}_n\}$ via the matrix \mathbf{U} . We will quantify the optimality of the sample positions in Section 2.5.

2.3. Sampling and reconstruction of random fields

In this Section, we introduce random fields that will be used to model and quantify spatial field correlations as well as to introduce ways to optimize sample locations. We express the field as $f(\vec{r}) = \boldsymbol{\psi}(\vec{r})^T \mathbf{a}$, where $\{\boldsymbol{\psi}_m\}$ can be any set of (not necessarily orthonormal) basis functions (e.g., u_m or f_i). We incorporate the prior knowledge by assigning a prior distribution for the coefficients a_m , which we consider random variables. We assume them to be Gaussian with a joint probability density $\mathcal{N}(\mathbf{m}_a, \mathbf{K}_a)$, where \mathbf{m}_a is the prior mean of \mathbf{a} and \mathbf{K}_a is the covariance matrix of \mathbf{a} .

A linear combination of basis functions with Gaussian random coefficients is a *Gaussian random field*, an extension of the Gaussian process (Williams and Rasmussen, 2006) to 3-D space. A random field $f(\vec{r}) = \boldsymbol{\psi}(\vec{r})^T \mathbf{a}$ can be described by its *mean field*

$$\mu_f(\vec{r}) = \mathbb{E}(f(\vec{r})) = \sum_m \mu_a[m] \boldsymbol{\psi}_m(\vec{r}) = \boldsymbol{\psi}(\vec{r})^T \boldsymbol{\mu}_a \quad (14)$$

and *covariance kernel*

$$\begin{aligned} K_f(\vec{r}, \vec{r}') &= \text{Cov}(f(\vec{r}), f(\vec{r}')) \\ &= \sum_{m, m'} \mathbf{K}_a[m, m'] \boldsymbol{\psi}_m(\vec{r}) \boldsymbol{\psi}_{m'}(\vec{r}') \\ &= \boldsymbol{\psi}(\vec{r})^T \mathbf{K}_a \boldsymbol{\psi}(\vec{r}'). \end{aligned} \quad (15)$$

Every finite collection of samples of a random field is jointly distributed as $\mathcal{N}(\boldsymbol{\mu}, \mathbf{K})$, where \mathbf{K} is the *sample covariance matrix* defined elementwise as $\mathbf{K}[n, n'] = K_f(\vec{r}_n, \vec{r}_{n'}) = \boldsymbol{\psi}(\vec{r}_n)^T \mathbf{K}_a \boldsymbol{\psi}(\vec{r}_{n'})$, and $\boldsymbol{\mu}$ is the sample mean, i.e., $\boldsymbol{\mu}[n] = \mu_f(\vec{r}_n) = \boldsymbol{\psi}(\vec{r}_n)^T \boldsymbol{\mu}_a$.

The effect of measurement is encoded in the posterior distribution of the field as illustrated in Fig. 2. Having a collection of (noisy) samples \mathbf{y} at sampling points $R = \{\vec{r}_n | n = 1, \dots, N\}$, modeling the noise as $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \boldsymbol{\Sigma})$, and assuming $\mathbb{E}(f(\vec{r})) = 0$, the posterior mean field and covariance are (Williams and Rasmussen, 2006)

$$\begin{aligned} \mu_f(\vec{r} | \mathbf{y}, R) &= \mathbf{k}(\vec{r})^T (\mathbf{K} + \boldsymbol{\Sigma})^{-1} \mathbf{y}, \\ K_f(\vec{r}, \vec{r}' | R) &= K_f(\vec{r}, \vec{r}') - \mathbf{k}(\vec{r})^T (\mathbf{K} + \boldsymbol{\Sigma})^{-1} \mathbf{k}(\vec{r}'), \end{aligned} \quad (16)$$

where $\mathbf{k}(\vec{r})$ is the covariance between the field at the sample points and the field at the point \vec{r} : $\mathbf{k}(\vec{r})[n] = K_f(\vec{r}, \vec{r}_n) = \boldsymbol{\psi}(\vec{r})^T \mathbf{K}_a \boldsymbol{\psi}(\vec{r}_n)$. Alterna-

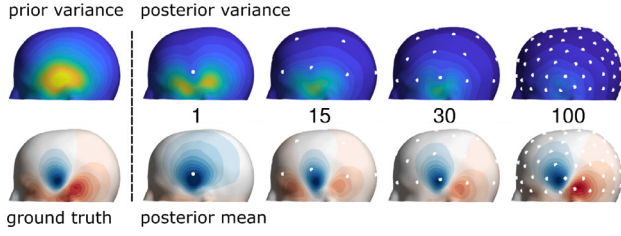


Fig. 2. Bayesian estimation of a random field. **Left:** Prior variance and a ground-truth field. **Right:** Posterior variance and the field reconstruction (the posterior mean field) after 1, 15, 30 and 100 noisy measurements (white dots) taken uniformly on the surface. Each additional sample decreases the posterior variance, yielding a better reconstruction of the field.

tively, the posterior mean and covariance of $f(\vec{r})$ can be expressed as [(Kailath et al., 2000) Sec 3.4]

$$\begin{aligned}\mu_f(\vec{r} | \mathbf{y}, R) &= \boldsymbol{\psi}(\vec{r})^T \mathbf{K}_a \boldsymbol{\Psi}^T \boldsymbol{\Sigma}^{-1} \mathbf{y} = \boldsymbol{\psi}(\vec{r})^T \hat{\mathbf{a}}, \\ K_f(\vec{r}, \vec{r}' | R) &= \boldsymbol{\psi}(\vec{r})^T \mathbf{K}_a \boldsymbol{\psi}(\vec{r}'),\end{aligned}\quad (17)$$

where $\hat{\mathbf{a}}$ is the posterior mean of the coefficients and $\mathbf{K}_a = (\boldsymbol{\Psi}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\Psi} + \mathbf{K}_a^{-1})^{-1}$ is the associated posterior covariance matrix.

The posterior mean $\mu_f(\vec{r} | \mathbf{y}, R)$ depends linearly on the samples \mathbf{y} and can be considered as the reconstruction of the field similar to Eq. (10). The posterior covariance $K_f(\vec{r}, \vec{r}' | R)$ describes the uncertainty in the field after the measurements; the noisier the samples, the more uncertainty is left. Uncertainty is reduced at all points \vec{r} that are correlated with the field at the sampling location.

2.4. Kernels for bioelectric potential and magnetic field

In this Section, we outline how to construct prior covariance kernels (Eq. (15)) for a random field on the measurement surface. In the later Sections, we discuss how these kernels can be used to obtain sampling grids.

Spatial-frequency kernel The SF basis functions can be used to express the prior kernel as $K(\vec{r}, \vec{r}') = \mathbf{u}(\vec{r})^T \mathbf{K}_{\text{SF}} \mathbf{u}(\vec{r}')$, where $\mathbf{u}(\vec{r})[i] = u_i(\vec{r})$ is the i th eigenfunction of the LB operator and \mathbf{K}_{SF} is the covariance matrix of the spatial-frequency coefficients. For more details of kernel representation with Laplacian eigenfunctions, see Solin and Särkkä (2020).

The assumption of a bandlimited field (Section 2.1) can be encoded by setting the diagonal of \mathbf{K}_{SF} to a constant up to B components and to zero above B ; the highest wavenumber of the field is assumed to be k_B . The covariance structure of this kernel is visualized in Fig. 3A both in the coefficient \mathbf{K}_{SF} and spatial domains $K(\vec{r}, \vec{r}')$. The spatial profiles $K(\vec{r}, \vec{r}_i)$ are sinc-like similar to reconstruction kernels in uniform 1D sampling of bandlimited functions (Jerri, 1977) and the spatial variance $K(\vec{r}, \vec{r})$ roughly uniform.

Dipole-field kernel Considering the source amplitudes q_i in Eq. (2) to be Gaussian random variables (de Munck et al., 1992), the covariance kernel is

$$K(\vec{r}, \vec{r}') = \mathbf{f}(\vec{r})^T \mathbf{K}_q \mathbf{f}(\vec{r}') = \sum_{i,j} \mathbf{K}_q[i, j] f_i(\vec{r}) f_j(\vec{r}'), \quad (18)$$

where \mathbf{K}_q is the $(M \times M)$ source covariance matrix. The dipole fields $f_i(\vec{r})$ can be considered as the basis functions that assemble the kernel and q_i as the associated (possibly correlated) random coefficients. Compared to the SF kernel, more detailed prior knowledge of the random field can be encoded in this form. For a set of N sampling points, the kernel reduces to a covariance matrix $\mathbf{K} = \mathbf{L} \mathbf{K}_q \mathbf{L}^T$, where \mathbf{L} is the discrete lead-field matrix.

A relevant kernel can be constructed by assuming the source amplitudes q_i identically and independently distributed (IID source prior): $\mathbf{K}_q = q^2 \mathbf{I}$. This kernel $K_{\text{IID}}(\vec{r}, \vec{r}') = q^2 \mathbf{f}(\vec{r})^T \mathbf{f}(\vec{r}')$ has non-uniform spatial variance and the spatial profiles $K(\vec{r}, \vec{r}_i)$ are asymmetric (Fig. 3A).

Kernel eigendecomposition and eigenbasis Principal component analysis can be extended to random processes using the Karhunen–Loève decomposition, which corresponds to finding the eigenfunctions of the covariance kernel (Loeve, 1978; Mercer, 1909). The Karhunen–Loève theorem (Stark and Woods 1986, section 7.6) states that the eigenbasis has the minimal truncation error among all possible bases for representing the random process with a given number of components. Any kernel can be written using its eigendecomposition as

$$K(\vec{r}, \vec{r}') = \mathbf{v}(\vec{r})^T \mathbf{D} \mathbf{v}(\vec{r}') = \sum_m d_m^2 v_m(\vec{r}) v_m(\vec{r}'), \quad (19)$$

where eigenfunctions $v_m(\vec{r})$ form an orthonormal basis on the measurement surface and \mathbf{D} is a diagonal matrix with variances d_m^2 of the eigenfunctions on the diagonal.

When studying the degrees of freedom in a random field generated by IID random sources, it is useful to express the total field variance as a sum of the eigenvalues of K_{IID} , i.e., $E(\|f\|^2) = \sum_{m=1}^{\infty} d_m^2$. With similar arguments as in Section 2.2, we can get an estimate for the number of samples needed to reconstruct the random field, e.g., by truncating the series up to the number of components $\{v_m\}$ that capture 99% of the total variance. A similar method has been applied before to study the sampling of EEG (Vaidyanathan and Buckley, 1997). Additionally, the number of components in the dipole fields f_i can be analyzed by projecting them on the eigenbasis of K_{IID} as done with the SF basis in Section 2.1. In Fig. 3B, we compare the SF basis and the eigenbasis of K_{IID} .

Noise kernels and SNR Measurement noise can be modeled as a random field $\mathbf{v}(\vec{r})$ with an associated covariance kernel $K_v(\vec{r}, \vec{r}')$. Sensor noise is typically independent across the sensors with an equal variance σ^2 , which can be modeled as an equivalent spatial white noise, i.e., a random field with a kernel $K_v(\vec{r}, \vec{r}') = \sigma^2 \delta(\vec{r} - \vec{r}')$. Projected to any orthonormal basis, the covariance of white noise is proportional to the identity matrix. Noise that originates from sources other than the sensors is generally colored (i.e., not white) and can be modeled with a covariance function $K_v(\vec{r}, \vec{r}') = \boldsymbol{\psi}_v(\vec{r})^T \mathbf{K}_v \boldsymbol{\psi}_v(\vec{r}')$, where $\boldsymbol{\psi}_v(\vec{r})$ are the basis functions for the noise field and \mathbf{K}_v is the prior covariance of the noise coefficients.

We can also define the signal-to-noise ratio (SNR) of a random field. We first diagonalize the noise kernel as in Eq. (19): $K_v(\vec{r}, \vec{r}') = \mathbf{w}(\vec{r})^T \boldsymbol{\Lambda} \mathbf{w}(\vec{r}')$, where $w_i(\vec{r})$ are eigenfunctions of K_v and $\boldsymbol{\Lambda}$ contains the associated noise variances on its diagonal. The noise eigenfunctions can be used to compose a *whitening kernel* $W(\vec{r}, \vec{r}') = \mathbf{w}(\vec{r})^T \boldsymbol{\Lambda}^{-1/2} \mathbf{w}(\vec{r}')$, which, when applied to the noise field as $\int W(\vec{r}, \vec{r}') \mathbf{v}(\vec{r}') dS'$, yields spatial white noise. Applying the whitener to a random field, we get a whitened field $\tilde{f}(\vec{r}) = \int W(\vec{r}, \vec{r}') f(\vec{r}') dS'$. The spatial SNR, i.e., the expected SNR for a sample at \vec{r} , is the variance of the whitened field

$$\text{SNR}(\vec{r}) = K_{\tilde{f}}(\vec{r}, \vec{r}) = \int \int W(\vec{r}, \vec{r}') K_f(\vec{r}', \vec{r}'') W(\vec{r}, \vec{r}'') dS' dS''. \quad (20)$$

With white noise, the whitening operation only acts as scaling with $1/\sigma^2$. With colored noise, it also modifies the covariance structure of the random field.

2.5. Optimality criteria of sampling grids

In this and the following Section, we discuss how an optimal sampling grid R can be constructed. A classical approach would be to choose R that minimizes the expected reconstruction error or the posterior variance, involving the inversion either of $\mathbf{U}^T \mathbf{U}$ (Section 2.2) or $\mathbf{K} + \boldsymbol{\Sigma}$ (Section 2.3). One way to obtain an optimal sampling is to find an R that maximizes the “size” of either of these matrices. The matrix size can be generally quantified in multiple ways, giving different optimality criteria in the context of *experimental design* (Chaloner and Verdinelli, 1995; Krause et al., 2008).

A common criterion is *A-optimality* (Krause et al., 2008) measured either as $\text{Tr}[(\mathbf{U}^T \mathbf{U})^{-1}]$ or more generally as $\text{Tr}[(\boldsymbol{\Psi}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\Psi} + \mathbf{K}_a^{-1})^{-1}]$. However, when studying sampling on a bounded surface as in MEG and EEG,

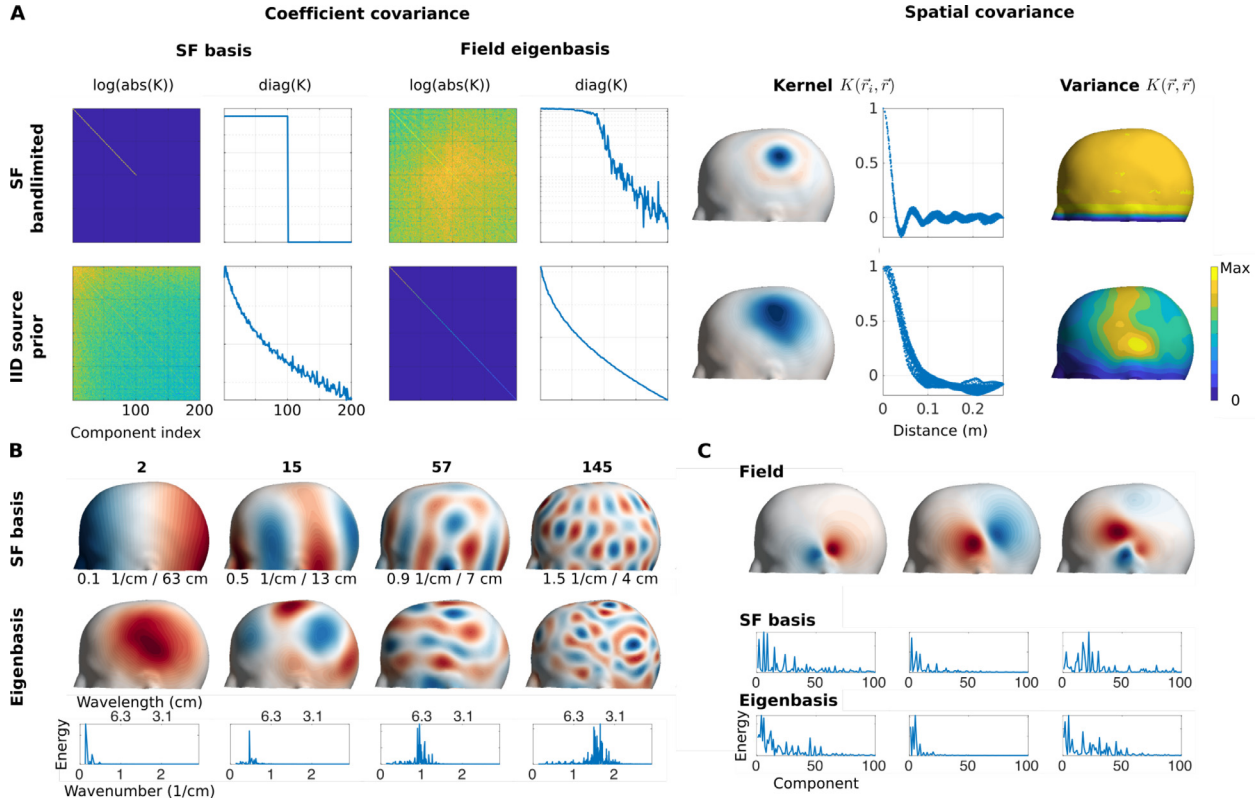


Fig. 3. A: Field covariances corresponding to two prior models: the field is bandlimited in spatial-frequency basis (SF bandlimited) and the field is generated by identically and independently distributed neural sources (IID source prior). **A, left:** Prior coefficient covariance in the SF basis and in the eigenbasis of K_{IID} . The covariance matrix and its diagonal are shown. **A, right:** Spatial covariance. The spatial kernel and its decay are shown for one position on the measurement surface. The rightmost column shows the spatial variance. **B:** Comparison of the spatial-frequency (SF) basis and the eigenbasis of K_{IID} . Each column shows the SF function and eigenfunction with the corresponding index. Wavenumbers and approximate wavelengths of the SF functions are displayed. The energy spectrum of the eigenfunction in SF basis is shown at bottom. The SF basis was generated using the zero-Neumann boundary condition. The eigenfunctions comprise multiple SF components with an average wavenumber proportional to the index. **C:** Example fields and their energy spectra both in the SF basis and eigenbasis. Energy is distributed to lower index components in the eigenbasis than in the SF basis indicating compression by the eigenbasis. For details on the computation, see Section 3.1. Units in the plots are arbitrary.

we should rather measure how well the sampling pattern captures the overall variance on the surface. To measure optimality in this sense, we define the fractional explained variance:

$$\text{FEV}(R) = 1 - \frac{\int K_f(\vec{r}, \vec{r}) |R| dS}{\int K_f(\vec{r}, \vec{r}) dS} = \frac{\int \mathbf{k}(\vec{r})^\top (\mathbf{K} + \mathbf{\Sigma})^{-1} \mathbf{k}(\vec{r}) dS}{\int K_f(\vec{r}, \vec{r}) dS}, \quad (21)$$

which ranges from 0 to 1, i.e., from no to all variance explained. This measure is related to *I-optimality* or integrated optimality (Atkinson, 2014).

The matrix size can also be measured using the determinant; this criterion is called *D-optimality* (Chaloner and Verdinelli, 1995). This leads to maximization of the total information (Appendix B)

$$\text{TI}(R) = \frac{1}{2} \log_2 \frac{\det(\mathbf{K} + \mathbf{\Sigma})}{\det(\mathbf{\Sigma})} = \frac{1}{2} \log_2 \det(\tilde{\mathbf{K}} + \mathbf{I}), \quad (22)$$

where $\tilde{\mathbf{K}} = \mathbf{\Sigma}^{-1/2} \mathbf{K} \mathbf{\Sigma}^{-1/2}$ is the whitened sample covariance matrix. The D-optimal R minimizes the posterior entropy of the random-field coefficients \mathbf{a} (Sebastiani and Wynn, 2000), i.e., maximizes the information gained, e.g., from the neural sources. Previously, total information has been used in studies comparing MEG sensor arrays (Iivanainen et al., 2017; Kempainen and Ilmoniemi, 1989; Nenonen et al., 2004; Riaz et al., 2017; Schneiderman, 2014).

2.6. Sampling grid construction

We now propose a method to obtain sampling grids that maximize the total information for given prior assumptions. In Appendix B, we

show that this corresponds to maximizing the diagonal elements in the whitened sample covariance matrix $\tilde{\mathbf{K}}$ while simultaneously minimizing the absolute values of the non-diagonal elements. In other words, maximizing total information is equivalent to finding the sampling grid with the least correlations and maximal SNR.

As an illustrating example, we first discuss how the sample spacing in the Shannon–Nyquist theorem (Jerri, 1977; Shannon, 1949) can be seen to result from information maximization. When treating bandlimited functions as random processes with a uniform prior variance for the spatial-frequency coefficients up to k_B , the kernel $K(x, x')$ can be calculated as the Fourier transformation of a boxcar function. This results in a sinc-function kernel with zeros at equispaced intervals $1/(2k_B)$ illustrated in Fig. 4. When the samples are placed at the zero crossings, the sample covariance matrix \mathbf{K} becomes diagonal, maximizing the total information.

For a general case, we reformulate the problem using the kernel eigenfunctions $v_n(\vec{r})$. We consider a (whitened) kernel $K(\vec{r}, \vec{r}')$ with an eigendecomposition $\mathbf{v}(\vec{r})^\top \mathbf{D} \mathbf{v}(\vec{r}')$. The decomposition can be rewritten as $K(\vec{r}, \vec{r}') = \tilde{\mathbf{v}}(\vec{r})^\top \tilde{\mathbf{v}}(\vec{r}')$, where $\tilde{\mathbf{v}}(\vec{r}) = \mathbf{D}^{1/2} \mathbf{v}(\vec{r})$, i.e., the covariance between points \vec{r} and \vec{r}' can be calculated as an Euclidean dot product between $\tilde{\mathbf{v}}(\vec{r})$ and $\tilde{\mathbf{v}}(\vec{r}')$. Rewriting the squared distance between these vectors as

$$\begin{aligned} \|\tilde{\mathbf{v}}(\vec{r}) - \tilde{\mathbf{v}}(\vec{r}')\|^2 &= \|\tilde{\mathbf{v}}(\vec{r})\|^2 + \|\tilde{\mathbf{v}}(\vec{r}')\|^2 - 2\tilde{\mathbf{v}}(\vec{r})^\top \tilde{\mathbf{v}}(\vec{r}') \\ &= K(\vec{r}, \vec{r}) + K(\vec{r}', \vec{r}') - 2K(\vec{r}, \vec{r}'), \end{aligned} \quad (23)$$

we can see that maximizing this distance with respect to sample positions \vec{r} and \vec{r}' corresponds to maximizing the diagonal elements in the

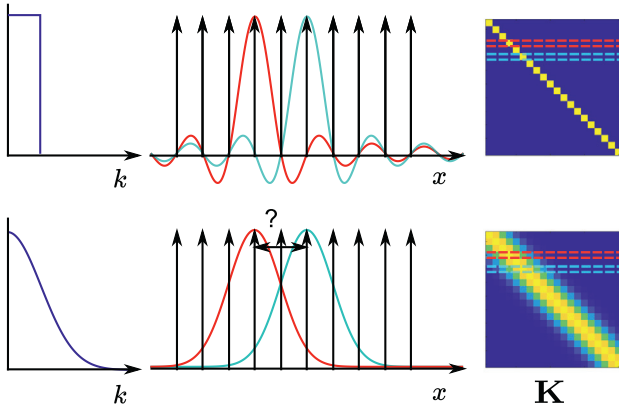


Fig. 4. Illustration of covariance kernel and sampling. **Left:** Coefficient variance as a function of spatial frequency for a strictly bandlimited process and a process with a smooth variance decay. **Middle:** The covariance functions and samples. The red and cyan curves describe the covariance functions for two samples while the spikes show the sample positions. Equispaced samples fit the zeros of the sinc functions. For a non-sinc covariance, there is no trivial choice for the optimal sampling distance. **Right:** Sample covariance matrices for the two cases, i.e., the covariance functions sampled at the equispaced locations. Dashed lines indicate the rows corresponding to the covariance functions on the left.

sample covariance matrix \mathbf{K} while simultaneously minimizing the non-diagonal elements.

The sample configuration that yields maximal information can be found with the *farthest-point sampling* algorithm (Eldar et al., 1997; Schlömer et al., 2011), which attempts to maximize the pair-wise minimum distances between the sample points. Instead of maximizing distances $|\vec{r} - \vec{r}'|$, we maximize $\|\vec{v}(\vec{r}) - \vec{v}(\vec{r}')\|$; otherwise the algorithm works similarly. If $\mathbf{K}[i, j] = K(\vec{r}_i, \vec{r}_j)$ are positive for neighbouring samples, the algorithm leads to minimization of the non-diagonal $|\mathbf{K}[i, j]|$, and thereby to maximization of the total information.

Generally, the sampling grid given by the method follows the covariance structure of the kernel (Fig. 3). For example, approximately uniform sampling grids can be generated using spatial-frequency-bandlimited covariance with circularly symmetric sinc-like kernels (Section 3.4). This showcases a connection between our sampling method and the sampling theorems that assume bandlimited functions and use uniform sampling (Pesenson, 2015; Shannon, 1949).

3. Simulations

In this Section, we simulate spatial-frequency (SF) spectra of dipole fields and covariance kernels of random fields due to random source distributions. Further, we construct optimal sampling grids for different kernels and evaluate their performances. We begin by outlining the numerical methods.

3.1. Models and computations

We computed the electric potential and magnetic field using linear Galerkin boundary-element method with the isolated-source approach (Stenroos et al., 2007; Stenroos and Sarvas, 2012). The head was assumed a piecewise constant isotropic conductor with four compartments: brain (the white and gray matter as well as the cerebellum), cerebrospinal fluid (CSF), skull and scalp. The conductivity of the soft tissues (brain and scalp) was set to 0.33 S/m, the conductivity of CSF to 1.79 S/m, and skull conductivity to 0.33/50 S/m. The head model was built from the example data of SimNIBS software (version 2.0; Windhoff et al. 2013) in an earlier study by Stenroos and Nummenmaa 2016. In addition to the conductivity interfaces the model included the grey-white-matter boundary.

We discretized the source-current distribution into point-like dipolar source-current elements. We placed the dipoles on the gray-white-matter boundary at 3-mm average spacing. The 20 324 sources were oriented normal to the surface following the anatomical orientation of apical dendrites of pyramidal neurons.

We calculated the electric potential and the normal component of the magnetic field at the nodes of triangular meshes that represented the measurement surfaces. The measurement surfaces were generated from a dense scalp mesh by cutting the mesh above a plane defined roughly by the ears and nose. The cut mesh was resampled to 5 404 nodes and 10 421 triangles. The mesh for the electric potential (EEG) was obtained by projecting the node positions onto the scalp. An 'on-scalp' MEG surface was generated by inflating the mesh 4.5 mm away from the scalp. A more distant 'off-scalp' MEG surface was obtained by further inflating the mesh and by smoothing it with the function `smoothsurf` in the `iso2mesh` MATLAB toolbox (Fang and Boas, 2009). The median distance of the nodes of the off-scalp MEG surface to the scalp was 2.4 cm (2.0–4.4 cm).

We further computed the lead-field matrices for 102 magnetometers of a commercial SQUID-MEG system (MEGIN Oy, Helsinki, Finland) and for a 60-channel custom version of the ANT waveguard™ EEG cap (ANT Neuro, Hengelo, The Netherlands). The sensor locations were derived from a measurement file of the MEGIN system and were manually coregistered to the used head geometry. 57 electrodes of the 60-channel EEG layout were employed in the measurement; we used those for the lead-field computation. We denote these arrays as SQUID102 and EEG57, respectively (Fig. 10C & D). The SQUID magnetometers were modeled with four integration points as described in the MNE software (Gramfort et al., 2014).

The SF basis vectors were computed by discretizing the LB operator to the triangle mesh in the weak form (Reuter et al., 2009). The discrete form of the eigenvalue Eq. (3) is

$$-\mathbf{C}\mathbf{u}_i = k_i^2 \mathbf{M}\mathbf{u}_i, \quad (24)$$

where \mathbf{u}_i contains the nodal values for the i th SF function, \mathbf{M} is a matrix that takes account the overlap in the piecewise-linear basis functions, and \mathbf{C} is the discrete LB operator. Matrices \mathbf{C} and \mathbf{M} were computed using MATLAB functions `cotmatrix` and `massmatrix` included in `gptoolbox` (Jacobson et al., 2018). The zero-Neumann boundary condition, which sets the outwards-facing derivative of u to zero, was used in the spectral energy analysis, while the zero-Dirichlet boundary condition, setting u to zero, was used in the grid construction. The energy spectra were calculated by discretizing the inner product in Eq. (4) as

$$\langle u_i, u_j \rangle = \int_S u_i(\vec{r}) u_j(\vec{r}) dS \approx \mathbf{u}_i^T \mathbf{M} \mathbf{u}_j. \quad (25)$$

The dipole-field kernel K_{IID} (Section 2.4) was constructed by assuming the sources on the cortically-constrained source space identically and independently distributed. The eigenbasis of K_{IID} was computed by discretizing the kernel on the surface and solving the discrete eigenvectors \mathbf{v}_i .

3.2. Energy spectra and decompositions of dipole fields

Methods. We quantified the dipole-field energy spectra to estimate the beneficial numbers of spatial samples as discussed in Section 2.2. Specifically, we analyzed the number of SF components needed to get 99% energy of each dipole field. We denote this number by $B_{99\%}$ and define the corresponding wavenumber as the 99% bandwidth, $k_{99\%}$. The sensor spacing in each modality corresponding to the largest $B_{99\%}$ values across the sources was estimated using uniform sampling grids (see Section 3.4). Further, in order to quantify compressibility of the dipole fields, we calculated energy spectra in the eigenbasis of K_{IID} ; we denote the number of eigencomponents to achieve 99% energy as $M_{99\%}$.

We inspected how many SF components of the dipole fields were above spatial white noise for a given SNR. We defined the dipole-field

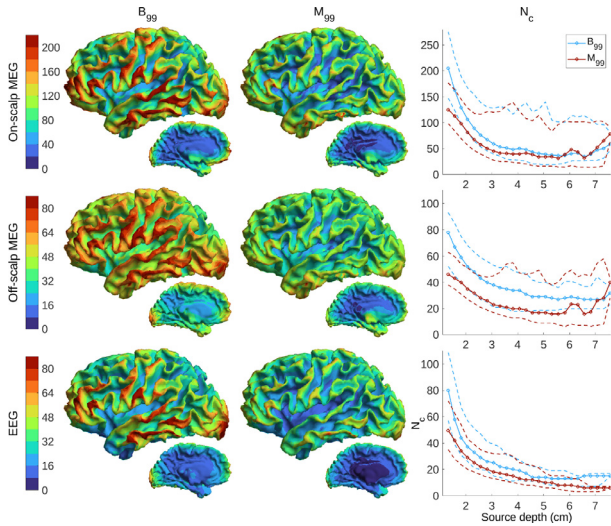


Fig. 5. The effective number of spatial components in dipole fields of on-scalp MEG, off-scalp MEG, and EEG for each source position on the cortical surface. **Left:** The number of spatial-frequency (SF) components to reach 99% of the energy ($B_{99\%}$). **Center:** The number of K_{IID} -eigenvectors to reach 99% energy ($M_{99\%}$). **Right:** $B_{99\%}$ and $M_{99\%}$ as function of source depth. The sources have been distributed to 2.5-mm-wide bins according to their depth. The solid lines plot the median values of $B_{99\%}$ and $M_{99\%}$ for the bins while the dashed lines indicate the 2.5 and 97.5 percentiles.

SNR as $\text{SNR}_f = \|f\|^2 / E(\|v\|^2)$, where $\|f\|^2$ is energy of the dipole field f and $E(\|v\|^2)$ is the expected energy of a sensor-noise-equivalent white-noise field v . For every modality and source, we set the ratio of dipole amplitude and white noise level so that the source had the given SNR. We then computed the number of SF components that had energy higher than the flat white noise spectrum. Three values of SNR were considered: 0.1, 1 and 10.

Results Fig. 5 displays $B_{99\%}$ and $M_{99\%}$ for sources on the left cortical hemisphere. Generally, $B_{99\%}$ decreases as a function of source depth. For the sources in the most superficial bin, the 97.5th percentile of $B_{99\%}$ is 276, 93 and 109 for on-scalp, off-scalp MEG and EEG, respectively. The bandwidths $k_{99\%}$ are then about 2.1, 1.0 and 1.3 1/cm, respectively, with corresponding uniform sensor spacings 1.6, 3.3 and 2.6 cm. Generally, $M_{99\%}$ is lower than $B_{99\%}$ indicating compressibility of the dipole fields, e.g., for the most superficial sources, the 97.5th percentiles of $M_{99\%}$ are 177, 63 and 72.

Field energies and numbers of SF components above noise are shown in Fig. 6. In MEG, the energy–depth curves have inverted V shape for the superficial sources as radial sources on gyri have smaller energies than tangential sources on sulcal walls. Additionally, MEG energy curves decay faster than those of EEG. For sources with SNR = 10, at maximum 205, 109 and 91 components are above noise in on-scalp/off-scalp MEG and EEG.

3.3. Field kernels

Methods We analyzed the spatial variance $K_{\text{IID}}(\vec{r}, \vec{r})$ (amount of signal) and correlations $K_{\text{IID}}(\vec{r}_i, \vec{r})$ (sample spacing; Section 2.4). The number of eigenvectors explaining 99% of the total variance was calculated to estimate the spatial degrees of freedom of the random source distribution (Section 2.4). Field correlation length was quantified for each measurement point \vec{r}_i by computing the distance along the surface at which $K_{\text{IID}}(\vec{r}_i, \vec{r})$ had decayed to half of its maximum value (the ‘half-maximum width’). The distances along the surfaces were computed using a method based on the heat equation (Crane et al., 2017) implemented as the MATLAB function `heat_geodesic` in `gptoolbox` (Jacobson et al., 2018).

Results Fig. 7 illustrates the kernels due to the IID random source distribution for the three modalities. To explain 99% of the total variance, 88, 35 and 26 eigenvectors are needed in on-scalp MEG, off-scalp MEG and EEG, respectively. Variance is distributed nonuniformly on the measurement surface with the highest values around the temporal cortex. The kernels are asymmetric and the correlation length varies across the surface; the least correlated area can be found on top of the temporal cortex. The correlation lengths are shortest in on-scalp MEG as indicated by smaller values of half-maximum widths compared to off-scalp MEG and EEG.

3.4. Sampling grid construction and evaluation

Methods The sampling grids were constructed by subsampling the nodal positions of the triangle meshes using the method described in Section 2.6. For a given kernel discretized on the mesh, we calculated its eigendecomposition, and utilized a farthest-point sampling algorithm implemented in `gptoolbox` (Jacobson et al., 2018) to obtain the optimal sampling positions.

Grid construction We constructed sampling grids for scenarios where random source distributions were defined in the brain: a global scenario where the whole brain was assumed active and of interest, and for a local scenario where a region of interest (ROI) in the brain was defined. We compared the performance of uniform sampling to model-informed sampling for different numbers of spatial samples N . Uniform sampling was generated using bandlimited priors in the SF basis (white covariance for the N lowest SF coefficients; SF bandlimited prior; Section 2.4). Model-informed sampling was generated by encoding the prior in the covariance of the cortically-constrained sources (IID source prior; Section 2.4). The grids were evaluated by computing total information (TI) and fractional explained variance FEV according to Eqs. (22) and (21) from the ‘ground-truth’ model of the signal and noise.

We ran the optimization algorithm several times with a random initial configuration as the output of the algorithm was sensitive to the initial condition. In the global scenario, the algorithm was run 20 times for $N < 25$, 10 times for $25 \leq N < 50$ and twice for $N \geq 50$. More iterations were chosen for small N due to the higher variance in TI. To seek a globally optimal grid with model-informed sampling, we took the grid with maximum TI for each N . For the uniform grids, we computed the average TI across the iterations as we wanted to quantify the general average performance of uniform sampling. In the local scenario, the algorithm was run once for each N ; N ranged in steps of 1 so that one iteration for each N was deemed sufficient to extract a general trend of TI as a function N .

3.4.1. Global scenario

Methods Here, the ground-truth model corresponded to the IID source prior and spatial white noise. The grids were constructed on meshes with the regions of eyes cropped out. We performed two analyses. In the first, we compared uniform and model-informed sampling in EEG and MEG as a function of SNR. The white noise variance ranged from $0.05\sigma_0^2$ to $20\sigma_0^2$. The reference noise level σ_0 was fixed in each modality so that it gave an average SNR (Eq. (20)) of 1 for source standard deviation of 7.6 pAm. Using this convention, σ_0 was determined to be 9 fT, 3.7 fT and 42 nV in on-scalp, off-scalp MEG and EEG, respectively. TI was also computed for MEG102 and EEG57; σ_0 was calculated to be 2.9 fT and 42 nV, respectively.

In the second analysis, we compared uniform sampling of on-scalp and off-scalp MEG as a function of spatial white noise level. We set the reference source variance q_0^2 so that the average SNR of off-scalp MEG was 1 with a white noise level of 3 fT. The white noise level of on-scalp MEG ranged from 1 to 20 fT and the source variance was varied ($0.2q_0^2$, q_0^2 and $5q_0^2$). On-scalp MEG grids comprising 10–600 points were generated, while off-scalp MEG had 100 or 300 points.

Results Fig. 8 summarizes the first analysis in the global scenario. For model-informed sampling, the sample density is the highest on the

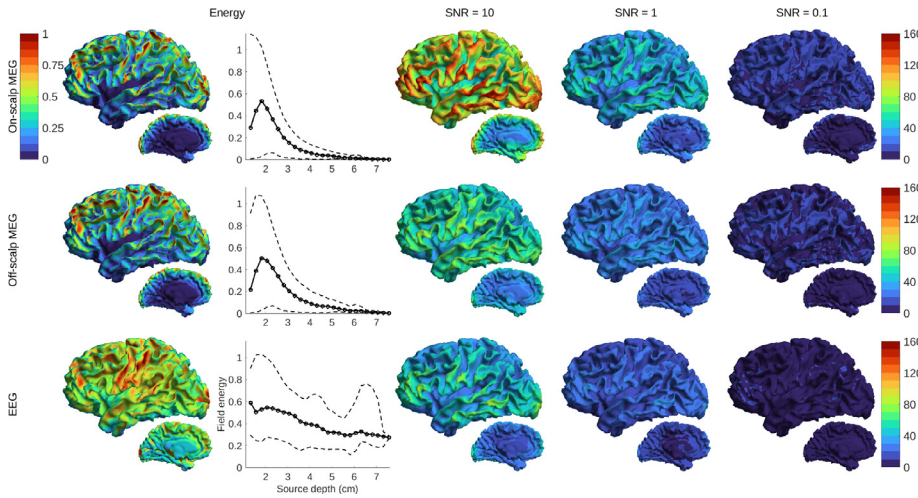


Fig. 6. Dipole-field energies and numbers of spatial-frequency (SF) components above noise. **Left:** Normalized dipole-field energies for on-scalp MEG, off-scalp MEG, and EEG for each source on the left cortical hemisphere and the same data as a function of source depth (the bins and the statistics are same as in Fig. 5). **Right:** The number of SF components in the dipole field above noise floor for each source position at different dipole-field SNR levels. Each dipole field has the same SNR, i.e., the source variance varies between the dipoles.

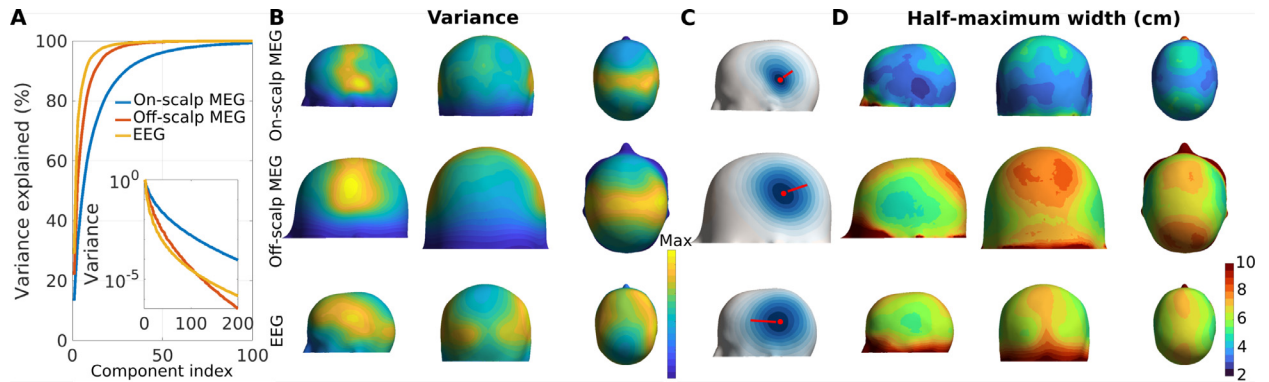


Fig. 7. Analysis of dipole-field kernel due to IID Gaussian sources. **A:** Variance as a function of the kernel eigenvector. Inset shows the normalized eigenvalues (variances). **B:** Distribution of the variance on the measurement surface. **C:** Spatial profile of the field kernel at one position (red dot) on the measurement surface and the distance at which the covariance has decayed to half (half-maximum width). **D:** Half-maximum width of the kernel across the measurement surface.

temporal lobe, which corresponds to the shorter correlation lengths and higher variance of the field kernel shown in previous analysis (Fig. 7). Model-informed sampling is especially beneficial in MEG at low SNR < 1 and small sample numbers, giving about 10–25% increase in TI compared to uniform sampling.

With the same SNR and the same number of samples, the TI is highest in on-scalp MEG among the modalities. When the SNR is low (0.1), approximately 60, 70 and 80% of variance is explained (on-scalp MEG, off-scalp MEG and EEG, respectively) with a sample number as high as 400. When the SNR is high (10), about 80, 30 and 20 samples are needed to explain 90% of the total variance.

Figure 9 gives a summary of the second analysis comparing on- and off-scalp MEG. With a similar noise level, SNR as averaged over the whole measurement surface is higher in on-scalp MEG by a factor of 6.0. Figure 9B shows TI and FEV as a function number of spatial samples and noise. With a comparable noise level, on-scalp MEG achieves the same information as 300 off-scalp samples with fewer samples.

Figure 9C gives the number of spatial samples needed in on-scalp MEG for the same TI as in off-scalp MEG as a function of noise level. For example, with a source variance of q_0^2 and noise levels of 3, 7, 10 and 15 fT in on-scalp MEG, about 30, 130, 260 and 590 samples are needed to yield the same TI as 300 off-scalp MEG samples at 3-fT noise level.

3.4.2. Local scenario

Methods In the local scenario, the ground-truth model consisted of IID sources distributed in an ROI defined around the motor cortex. Both white (sensor) and colored (sensor + background brain activity) noise

were considered. The ROI consisted of sources within a patch that had a radius of approximately 3 cm along the cortical surface (387 source elements; Fig. 10A). Sampling grids were constructed for on-scalp MEG and EEG. The total information was analyzed and SQUID102 and EEG57 were used as references.

The source variance q^2 was set so that the maximum SNR of the on-scalp MEG was 3 with a spatial white noise level of 9 fT. The white noise level of EEG (25 nV) was set so that the maximum SNR was also 3. To model colored noise due to brain background activity, the sources outside the ROI were assumed active (IID Gaussian; variance $q^2/100$). The noise level of SQUID102 was set to 3 fT while that of EEG57 was 24 nV.

The field kernel was whitened as described in Section 2.4 and SNR was analyzed. The eigenvalues λ_i of the whitened kernel were converted to bits as $1/2 \log_2(\lambda_i + 1)$. These values were used to analyze the number of eigenvectors contributing to the total information. Two uniform sampling schemes were considered. In the first, samples were distributed evenly on the whole head. In the second, the SF basis was constrained to a region on the surface that had a distance to the center of the ROI less than 9 cm. For model-informed sampling, a dipole-field kernel corresponding to the ground-truth model was used, while the noise kernel corresponded either to the white or colored noise model.

Results Fig. 10A shows the SNR distributions as well as the TI content across the eigenvectors. With white noise, the TI is 28 bits in MEG and 27 bits in EEG. With colored noise, the maximum SNR is 1.1 in MEG and 0.3 in EEG while the TI is 18 and 12 bits, respectively. The TI and maximum SNR in SQUID102 are 4.6 bits and 1.1 with white noise and

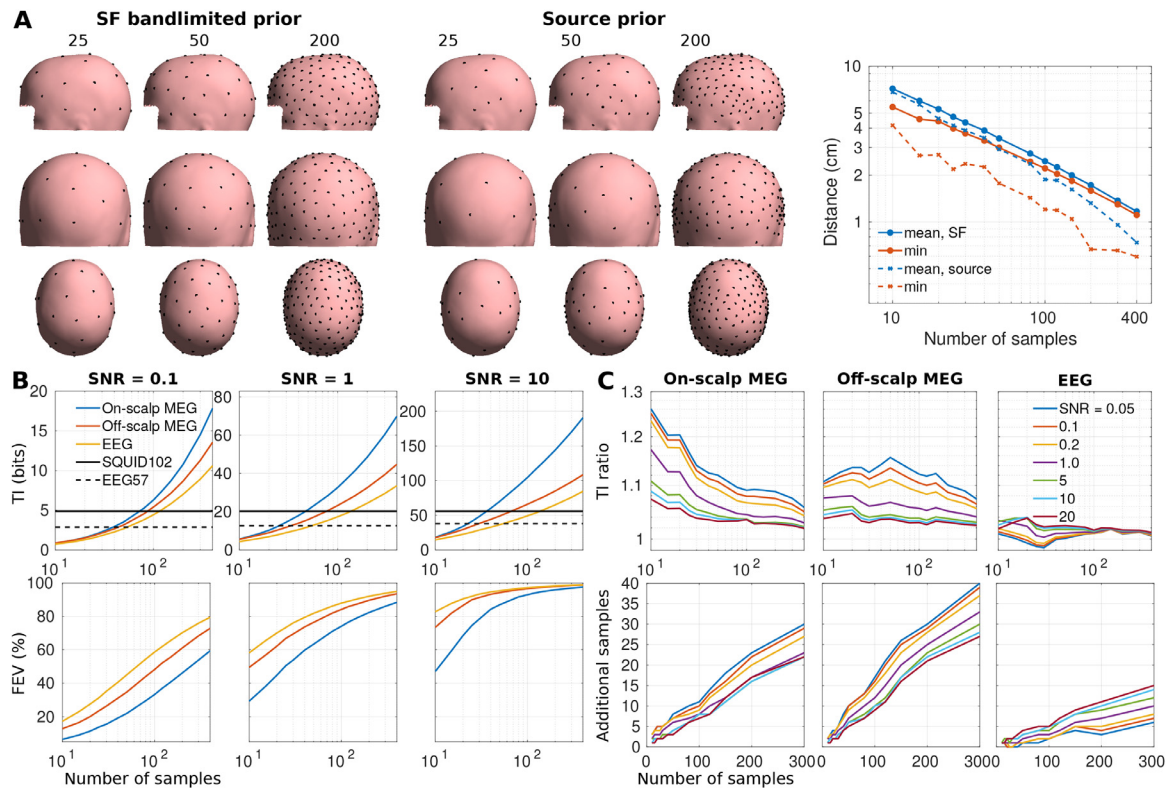


Fig. 8. Comparison of uniform (SF bandlimited prior) and model-informed (IID source prior) sampling of the whole brain in on-scalp MEG, off-scalp MEG and EEG. **A:** Example on-scalp MEG sampling grids generated with the priors and their average and minimum sample spacing. **B:** Total information (TI) and fractional explained variance (FEV) of the uniform grids. TI for SQUID102 and EEG57 are also shown. **C:** Ratio of the TI obtained by model-informed and uniform sampling with different SNRs (top) and the number of additional samples required with uniform sampling to achieve the same TI as with model-informed sampling (bottom).

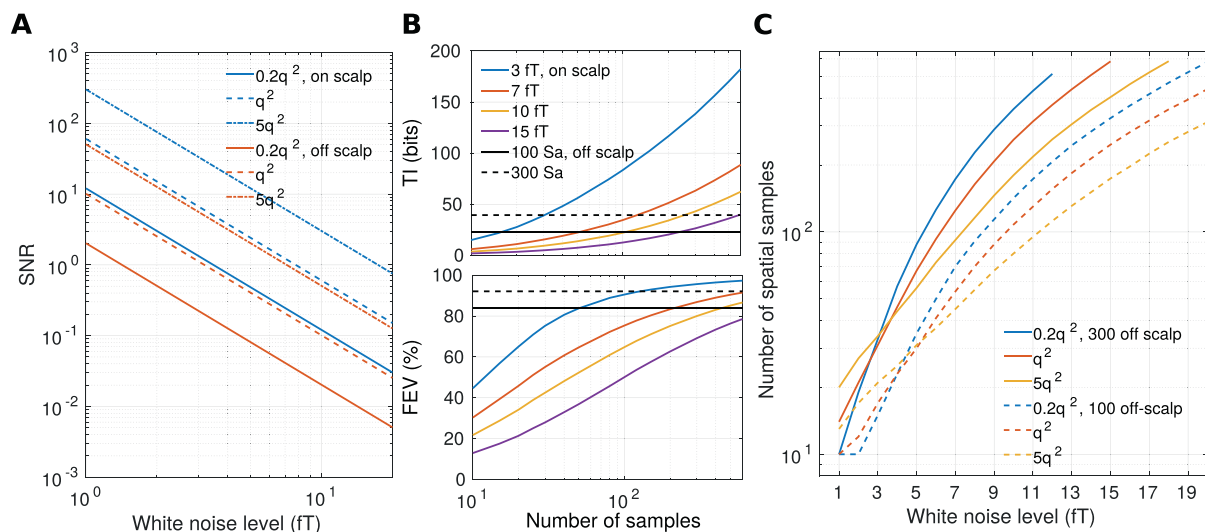


Fig. 9. Spatial sampling of independent sources (variance q_0^2) distributed over the whole cortex with on-scalp and off-scalp MEG. **A:** Signal-to-noise ratio averaged over the whole measurement surface as a function of spatial white noise level. **B:** Total information (TI) and fractional explained variance (FEV). Colored lines give the values for on-scalp MEG as a function of sample number and noise level. The vertical lines show the values for off-scalp MEG with 100 and 300 samples with a noise level of 3 fT. The source variance is q_0^2 . **C:** Number of spatial samples needed in on-scalp MEG to achieve the same TI as in off-scalp MEG with a noise level of 3 fT as a function of on-scalp MEG white noise level. Comparisons to the TI of 100 off-scalp samples is plotted with the dashed line and to 300 samples with the solid line.

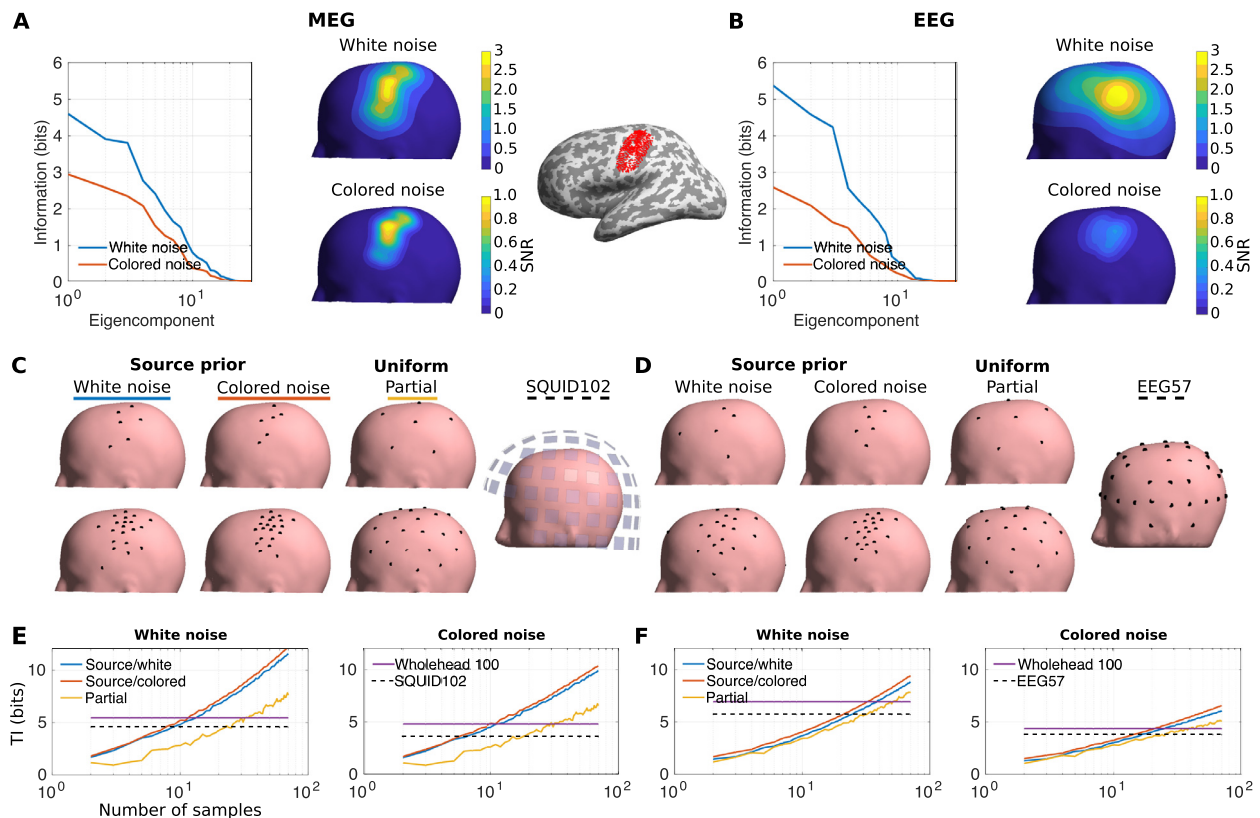


Fig. 10. Spatial sampling of on-scalp MEG (left panel; A, C and E) and EEG (right; B, D and F) fields due to uncorrelated sources in an ROI around the motor cortex (shown on the center in red on the inflated cortical surface). **A & B:** Distribution of total information (TI) among the eigenvectors and the spatial distribution of SNR computed with spatial white noise as well as with colored noise (brain background activity + white noise). **C & D:** Example grids constructed with IID source and bandlimited SF (~uniform sampling) priors. With IID source prior, two noise priors (white and colored) were considered. Sensor layouts of SQUID102 and EEG57 arrays are also shown. **E & F:** TI as a function of number of samples in the grids computed using the two noise models. Wholehead 100 refers to a uniform wholehead grid with 100 spatial samples.

3.6 bits and 0.8 with colored noise, respectively. The corresponding TIs in EEG57 are 5.7 and 3.8 bits, respectively; the maximum SNRs are 2.4 and 1.3.

The grids constructed using the different priors for signal and noise are presented in Fig. 10B. Compared to uniform sampling, the model-informed grids are more densely distributed, especially when the colored noise model is used. Compared to whole-head sampling, using the source prior to distribute samples on the region with high SNR is beneficial. With colored noise, 10 samples are needed in the model-informed MEG grids to reach the same TI as with 100 whole-head samples. In EEG, 20 samples are needed.

4. Discussion

We quantified the number of spatial samples beneficial for MEG and EEG by analyzing the spatial-frequency content of fields due to dipolar sources. Based on a review of Gaussian processes and optimal design, we related the information metric used in MEG sensor-array designs to the covariance of random-field models. This relationship was used to derive an information-maximizing sampling algorithm. We applied the algorithm to generate sampling grids, which we used to quantify the benefit of model-informed sampling of random source distributions in comparison to uniform sampling.

4.1. Field analysis

To capture 99% of the field energy of every source in the brain of a representative adult male, approximately 280, 90 and 110 SF com-

ponents were needed in on-scalp, off-scalp MEG and EEG, respectively. These numbers represent the maximum number of spatial degrees of freedom in any field due to neural activity in the used head model. Additionally, they correspond to the number of uniform samples needed to achieve at most roughly 1% field reconstruction error in noiseless conditions. These results agree with previously published results. For MEG, the numbers corresponds to the "rule of thumb" presented by Ahonen et al. (1993): the sensor spacing should be approximately the distance of the sensors to the closest source. For EEG, the results are in line with the arguments by Srinivasan et al. (1998): the sensor count should be at least roughly 120 for adequate sampling of the EEG in a typical adult head.

The eigenbasis of the dipole-field kernel due to IID sources was found out to reduce the component number needed to capture 99% of the dipole-field energies, compressing the field representations. This compression and the variance analysis of IID sources suggest that the sample numbers given by the SF analysis provide oversampling, and, thereby, give a good upper limit on the beneficial sample number.

4.2. Head models

We employed a four-compartment realistically-shaped piecewise homogeneous head model. To our knowledge, previous sampling analyses have been carried out in simpler models such as half-infinite and spherically symmetric conductors. An essential property of the four-compartment model is the inclusion of geometrically complex gyral folding that strongly deviates from simple layer structure represented by spherical models or three-shell (3S) models commonly used

in EEG source analysis. Compared to EEG, MEG is only weakly sensitive to inclusion or omission of thin layer-like conductivity structures (Hamalainen and Sarvas, 1989; Stenroos et al., 2014). The inclusion of CSF, however, breaks the simple layer structure: the omission of CSF has roughly as large effects on MEG and EEG dipole fields, when compared with relative-error or correlation metrics (Stenroos and Nummenmaa, 2016).

While detailed analysis of different head models is beyond the scope of this manuscript, we did a simple model comparison to see, how the spatial frequency content is affected by the CSF and the skull, or the omission of them in the models. The comparison, presented in Appendix C, shows that the effective number of spatial-frequency components and the overall spatial spectrum in on- and off-scalp MEG are negligibly affected by the CSF. For EEG, we replicated the well-known low-pass filter property of a brain-skull-scalp system (e.g., Srinivasan et al. 1996), or more generally, of a layered structure of high, low and high conductivities. The highly conductive CSF enhances the low-pass phenomenon. The effect of the CSF on spatial-frequency decay, can, however, be partially compensated by lowering the conductivity of the skull in a 3S model, as was already demonstrated for EEG dipole fields in (Stenroos and Nummenmaa, 2016). The remaining difference in the energy spectrum is mainly of constant-gain nature. Most of the smearing of EEG is thus characterized by the parameters of the layer-like low-pass filter system, and the complex folding of CSF plays a smaller role.

Our conductor model did not contain separate white and gray matter or white-matter anisotropy, which affect EEG and MEG dipole fields (Guellmar et al., 2010; Vorwerk et al., 2014; Wolters et al., 2006). In the separation of white and gray matter in a model that contains the CSF, the amount of added conductivity contrast and deviation from layer structure are smaller than in the inclusion of CSF. The separation of white and gray matter should thus have a smaller effect on MEG and EEG energy decays and component counts than the CSF has. The white-matter anisotropy may have a non-negligible effect for sampling deeper sources with EEG.

Finally, our simulation was based on normal adult head. Strong deviations from layer structure between the source and sensors, such as fontanels in small children (Lew et al., 2013), surgical holes in the skull, or CSF-filled brain lesions, may have a strong effect on spatial information especially in EEG but also in MEG. These effects are a potential topic for future studies.

4.3. Grid construction

To generate sampling grids, we presented a method that utilizes prior information of signal and noise in the form of a covariance kernel. Assuming random spatial-frequency coefficients with a uniform variance up to some bandlimit, the kernel is isotropic and translation-invariant and the method yields uniform sampling grids similar to those in the sampling theorems introduced in Section 2.1. With this approach, uniform grids may be constructed also on more complex surfaces and domains. The kernel constructed from the dipole fields of IID random neural sources has a location-dependent shape and width, and the method yields nonuniform sampling grids.

In the whole-head scenario, nonuniform model-informed sampling was only slightly beneficial compared to uniform sampling in terms of total information. When the number of samples was small and the noise level high, nonuniform sampling yielded roughly 10–25% more information than uniform sampling in MEG. When decreasing the noise level or increasing the number of samples, the performance of uniform and nonuniform sampling grids became roughly similar. With nonuniform sampling, the same total information was, however, reached with fewer number of samples. In EEG, the benefits of nonuniform sampling were less clear; EEG may not benefit from it due to long-range correlations.

When a cortical region of interest was defined, local high-density (nonuniform) spatial sampling was beneficial; dense sampling of the lo-

cal components yielded higher total information than uniform sampling covering a larger surface area or the whole head. Dense sampling may be especially useful when colored noise fields (such as those generated by background brain activity) are present and the region of interest is small. Local dense arrays could be beneficial in applications where a certain part of the brain is of interest and the number of sensors is limited (e.g., Iivanainen et al., 2020) or in brain-computer interfacing where simple measurement setups are desirable.

4.4. Practical considerations and future directions

Our analysis corroborates the observation that the spatial-frequency content of MEG mostly depends on the source-to-sensor distance (Ahonen et al., 1993). Accordingly, for designing whole-scalp sensor arrays, the rule of thumb (sensor spacing approximately the distance of the sensors to the closest source) should give the approximate uniform sensor spacing. For example, in the case of an infant head where the source-to-sensor distance can be as small as 5 mm, this rule of thumb would suggest a sensor spacing of 5 mm. In addition to the spacing, the sensor array should have adequate coverage of the scalp so that most of the energy of the topographies can be captured by the spatial samples. For example, in the case of SQUID helmets, the limited coverage of frontal sources often leads to a loss of signal as well as spatial information from the frontal regions.

Our quantitative results are based on the analysis of a single adult head model and, thereby, cannot be directly generalized as there may be considerable variation in head sizes and shapes. For example, as mentioned earlier, smaller sensor spacing would be beneficial when measuring an infant. Thus, it would be valuable to perform similar computations in different subject populations and extract statistics on the metrics introduced in this paper. We note that there are freely available head-model collections that could be used for this purpose (see, e.g., Htet et al. 2019).

We demonstrated the developed method using a single ROI defined around the motor cortex. In future studies, the locations, sizes and number of ROIs could be varied and correlations between the sources could be included. The ROI could be defined to match the cortical area of experimental interest (e.g., V1 of the visual cortex).

The sampling grid optimization presented here does not necessarily result in sensor arrays that can be readily implemented in practice. For example, the minimum distance between the sensors was not constrained and the sensor dimensions were ignored. Sensor dimensions limit the minimum distance between the sensors and the sensor noise level is proportional to them in many sensor types (e.g., Kempainen and Ilmoniemi 1989; Mitchell and Alvarez 2020). Spatial integration within a sensor can be viewed as spatial low-pass filtering, an effect that can be used to reduce aliasing due to high spatial frequencies (Roth et al., 1989). These matters should be taken into account in future studies when designing practical sensor arrays. Also, the method could be used with a population of head models to design arrays with the best average performance.

We did not include in our simulation the effects of systematic or random sensor-positioning errors. Intuitively, the higher the spatial frequency, the more it will be affected by a given amount of sensor position error. By the same argument, sensor positioning errors might negate the benefits gained by using optimized sensor arrays. However, the sensor positioning error could be included in the optimization to some degree by blurring the kernel. We note that the spatial-frequency representation gives convenient means to assess how the spatial components of the field are affected by the sensor-position errors.

We analyzed scalar fields which is sufficient in EEG. In contrast, magnetic field is a vector field, and the orthogonal field components may each provide additional information about the neural sources (Iivanainen et al., 2017). The sensor orientation could be included in the optimization by applying vector basis functions to assemble the kernel. Also (external) interference fields could be taken into account. In this

case, vector-spherical harmonics could provide a useful basis (Taulu and Kajola, 2005). Sampling grids with triaxial sensors in a single layer (Brookes et al., 2021) or as multiple layers at different distances from the scalp should be beneficial in separating signals from the external interference (Nurminen et al., 2013; 2010). In fact, OPMs can provide dual- or tri-axial measurements of the field, making realistic implementations of such arrays feasible (Brookes et al., 2021). The methods presented here could be further developed and used to guide the design of such arrays. We note that the choice between a single- or multi-axis measurement for each sensor might not be trivial as the sensitivity of the OPM typically decreases when multiple components of the field are measured simultaneously (see, e.g., Osborne et al. 2018).

In MEG, it is common to use gradiometers instead of magnetometers. According to Ahonen et al. 1993, gradiometers that sample the tangential gradient of the normal field component can be placed at larger spacing than magnetometers. On the other hand, long-baseline axial gradiometers have similar lead fields as magnetometers (Malmivuo and Plonsey, 1995) suggesting that the results for those would be similar to what we obtained in this paper. However, adding external interference fields to the stochastic model would probably result in differences between the optimal magnetometer and axial-gradiometer arrays.

With the generalization of the SF basis to an arbitrary measurement surface, the standard 1-D signal processing methods such as filtering and interpolation should be straightforward to apply to spatial EEG/MEG data. Spatial-frequency filtering could be useful in noise and interference rejection as demonstrated by Graichen et al. (2015); the continuous SF basis functions make the filtering less dependent on the sensor configuration. As the SF basis is both data- and model-independent, the filtering methods may be useful in real-time applications. The SF filtering as a part of preprocessing would be similar to the spline Laplacian (Nunez et al., 2019), providing flexibility for defining the filter coefficients. Filtering and interpolation may also help in data visualization.

Estimating the neural sources that could have generated the data, i.e., solving the bioelectromagnetic inverse problem (Sarvas, 1987) is often of interest. We did not explicitly consider the inverse problem. However, the total information should be an estimator of the source-estimation performance as it quantifies the size (or the stability of the inversion) of the modelled or measured measurement covariance matrix. In other terms, when the kernel is generated using dipole fields, total information maximization corresponds to minimization of the posterior entropy of the sources (Section 2.5).

We conclude the discussion by summarizing the results from a practical perspective. Considering whole-scalp sampling, our results show that when the sensor count is sufficiently large (hundreds) the benefits from optimized sampling are minimal compared to uniform sampling. With lower sensor counts (tens), sampling that is optimized either for an individual subject or across a subject population might, however, be beneficial. On the other hand, considering the sampling of an a-priori determined cortical region, our results show that the scalp area with the strongest signal is quite compact and, thereby, dense sampling might be beneficial. OPMs that use a single vapor cell with multiple densely-packed channels (e.g., Xia et al. 2006) could be used to assess the benefits of dense sampling over a conventional array, e.g., in a classification task (for such an analysis with EEG, see Robinson et al. 2017).

5. Conclusions

We analyzed the spatial sampling of EEG and MEG and suggested a method for designing optimal sampling positions. Our simulations suggest that when measuring adult males on-scalp MEG can benefit from up to roughly 300 spatial samples while for off-scalp MEG and EEG this number is around 100. The theoretical framework we present in this manuscript as well as the sample-positioning method we introduce can be used to design sampling grids that convey the most information from the neuronal sources. Such designs may be useful when the sensor number is limited or a certain region of the brain is of interest.

Data and code availability statement

To facilitate further use in the research community, the information-maximizing sampling algorithm is available from the corresponding authors for academic use.

The data and code sharing policy adopted by the authors complies with the requirements of Aalto University.

Declaration of Competing Interest

The authors declare that they have no conflict of interest.

Credit authorship contribution statement

Joonas Iivanainen: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Software, Writing – original draft, Writing – review & editing. **Antti J. Mäkinen:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Software, Writing – original draft, Writing – review & editing. **Rasmus Zetter:** Conceptualization, Writing – review & editing. **Matti Stenroos:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Supervision, Writing – review & editing. **Risto J. Ilmoniemi:** Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing. **Lauri Parkkonen:** Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing.

Acknowledgments

This work has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 686865 (project BREAKBEN), the European Research Council under grant agreement No 678578 (project HRMEG), the National Institute of Neurological Disorders and Stroke of the National Institutes of Health under Award Number R01NS094604, the Finnish Cultural Foundation under grant Nos. 00170330 and 00180388 (JI), and Vilho, Yrjö and Kalle Väisälä Foundation (AM). The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding organizations.

Appendix A. Effective bandlimit

To determine an effective bandlimit for field f given by Eq. (2), we note that the spectral coefficients can be written as $a_m = \sum_i q_i \langle f_i, u_m \rangle$. Further assuming that the source amplitudes q_i are independent random variables with zero mean ($E(q_i q_j) = 0$), the expected energy of f can be decomposed using the spectra of the source fields:

$$E(\|f\|^2) = \sum_{m=1}^{\infty} \sum_i E(q_i^2) \langle f_i, u_m \rangle^2. \quad (\text{A.1})$$

If f_B in the bandlimited representation of f is chosen such that every source field f_i is expressed up to $1 - \alpha$ of its total energy (e.g., $1 - 0.01 = 0.99$), the expected energy of the residual field (the truncation error) can be bounded using Eq. (A.1) as

$$\begin{aligned} E(\|f_r\|^2) &= \sum_i E(f_i^2) \sum_{m=B+1}^{\infty} \langle f_i, u_m \rangle^2 \\ &\leq \sum_i E(f_i^2) \alpha \|f_i\|^2 = \alpha E(\|f\|^2). \end{aligned} \quad (\text{A.2})$$

Thus, any field due to independent neural sources is expected to have $1 - \alpha$ of its energy in the subspace spanned by the B first SF basis functions.

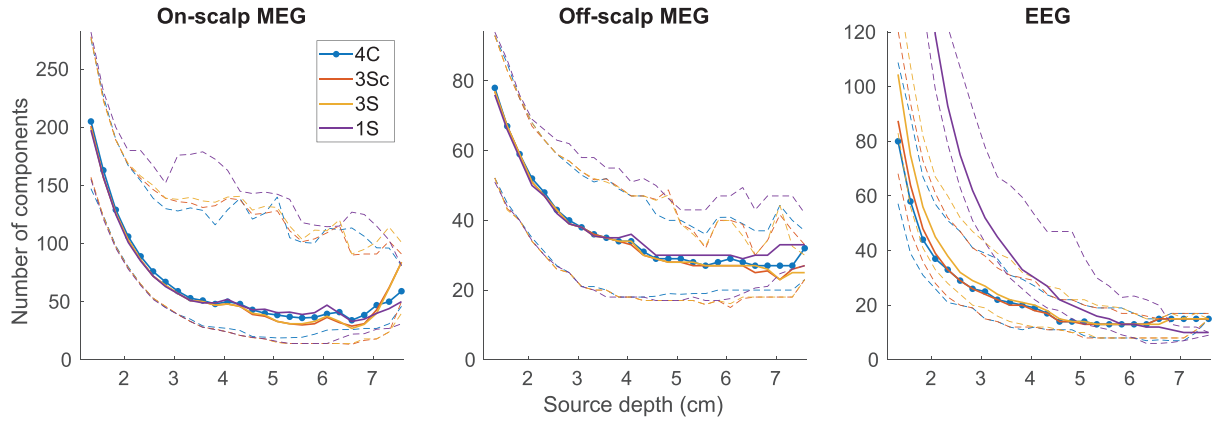


Fig. C.11. The effective number of spatial-frequency (SF) components for dipole fields ($B_{99\%}$) in different head models. The sources have been distributed to 2.5-mm-wide bins according to their depth. The solid lines with dots show the median values per bin, the dashed lines indicate the 2.5 and 97.5 percentiles.

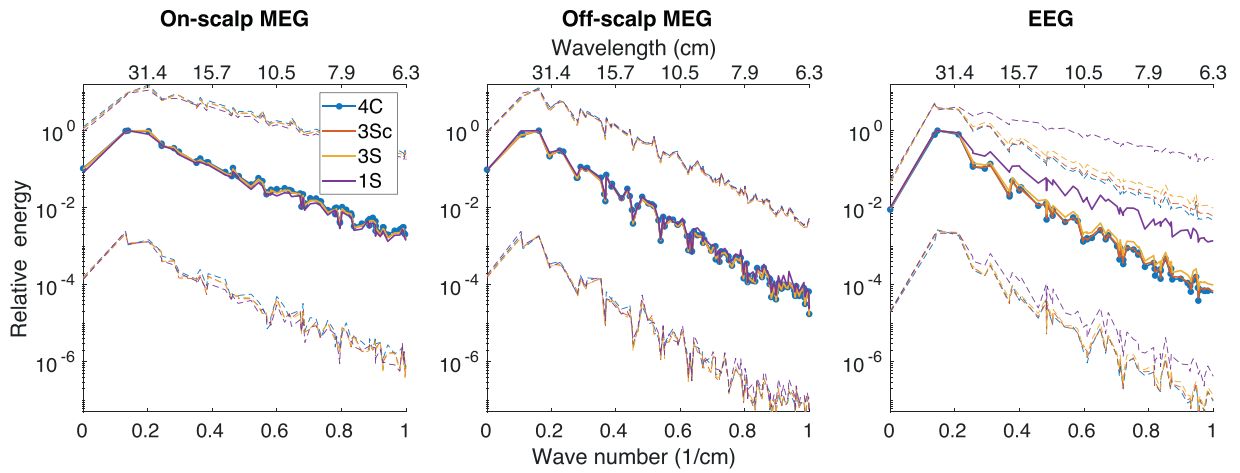


Fig. C.12. The spatial-frequency spectra of dipole fields in different models. The solid lines with dots show the median value across the cortex, the dashed lines indicate the 2.5 and 97.5 percentiles. The data of each model has been normalized to have unit maximum value in the median plot.

Appendix B. Total information and covariance

Here, we show how the total information (TI) relates to the sample and noise covariance matrices, \mathbf{K} and $\mathbf{\Sigma}$ (Section 2.3). If the samples as well as the noise were uncorrelated, TI would be $1/2 \sum_i \log_2(P_i + 1)$ where P_i is the power signal-to-noise ratio of each measurement. If the measurements are correlated, they can be orthogonalized by the eigen-decomposition of the whitened covariance matrix $\tilde{\mathbf{K}} = \mathbf{\Sigma}^{-1/2} \mathbf{K} \mathbf{\Sigma}^{-1/2} = \mathbf{V} \mathbf{P} \mathbf{V}^T$, where \mathbf{V} contains eigenvectors of $\tilde{\mathbf{K}}$ and \mathbf{P} is a diagonal matrix of P_i . Starting from the original formula, TI can be now written as

$$\begin{aligned} \text{TI}(R) &= \frac{1}{2} \sum_i \log_2(P_i + 1) = \frac{1}{2} \log_2 \det(\mathbf{P} + \mathbf{I}) \\ &= \frac{1}{2} \log_2 \det(\tilde{\mathbf{K}} + \mathbf{I}) = \frac{1}{2} \log_2 \frac{\det(\mathbf{K} + \mathbf{\Sigma})}{\det(\mathbf{\Sigma})}. \end{aligned} \quad (\text{B.1})$$

Total information can be maximized by choosing the measurement grid so that $\tilde{\mathbf{K}}$ is diagonal, i.e., each sample measures independent information. This can be seen, e.g., from the matrix derivative (Petersen and Pedersen, 2012)

$$\frac{\partial \ln \det(\mathbf{X} + \mathbf{I})}{\partial \mathbf{X}} = 2(\mathbf{X} + \mathbf{I})^{-1} - \mathbf{I} \odot (\mathbf{X} + \mathbf{I})^{-1}, \quad (\text{B.2})$$

where $\mathbf{I} \odot$ is element-wise product with identity matrix resulting in the diagonal-part of the matrix. The logarithm of determinant is maximized when the derivative with respect to its elements is zero, which means that the non-diagonal elements of $(\mathbf{X} + \mathbf{I})^{-1}$ must be zero. This is equiva-

lent to \mathbf{X} being diagonal itself. For diagonal elements, no solution exists, but their derivatives are always positive when \mathbf{X} is positive definite.

Appendix C. Head model comparison

To assess how different head tissues affect the spatial frequency content of measurements, we compared three simplified models to the four-compartment (4C) model. The roles of CSF compartment and skull conductivity were evaluated using two three-shell models with homogeneous compartments for the scalp, skull, and intracranial space. The conventional three-shell (3S) model had the same conductivities for the brain, skull and scalp as the 4C model. In the compensated three-shell model (3Sc), the conductivity of the skull was decreased to 0.33/80 S/m to compensate for the omission of the CSF (Stenroos and Nummenmaa, 2016). To assess the degree of lowpass filtering due to the brain-skull-scalp system on EEG and the role of skull in MEG, we used a single-shell (1S) model, bounded by the scalp.

We first compared $B_{99\%}$, the effective number of spatial components in dipole fields. The results are displayed in Fig. C.11. In MEG, the number of components is minimally affected by the omission of CSF, while the 1S model slightly exaggerates the level of detail with deeper sources. In EEG, the differences are largest with superficial sources; the 3S model somewhat exaggerates the number of components, while the 1S model produces essentially larger numbers, the median reaching 291 (vs. 80 in the 4C model) for the most superficial sources.

In addition, we compared spatial spectra for all sources, i.e., the contribution of each spatial-frequency component to the signal energy. The results are shown in Fig. C.12. For MEG, the spectra are nearly identical. In on-scalp MEG, the simplified models very slightly exaggerate the most superficial sources, which shows here as slightly larger values for the 4C model at deeper sources. For EEG, the 4C and 3-shell models produce very similar spectra, while the 1S model produces considerably larger values for high spatial frequencies.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at [10.1016/j.neuroimage.2021.118747](https://doi.org/10.1016/j.neuroimage.2021.118747).

References

- Abrahamsen, P., 1997. A review of Gaussian random fields and correlation functions. *Neurosci. Biomed. Eng.* 40 (9), 859–869.
- Ahonen, A.I., Hämäläinen, M.S., Ilmoniemi, R.J., Kajola, M., Knuutila, J.E.T., Simola, J., Vilkmann, V.A., 1993. Sampling theory for neuromagnetic detector arrays. *IEEE Trans. Biomed. Eng.* 40 (9), 859–869.
- Atkinson, A.C., 2014. Optimal Design. In: Wiley StatsRef: Statistics Reference Online, pp. 1–17.
- Baillet, S., 2017. Magnetoencephalography for brain electrophysiology and imaging. *Nat. Neurosci.* 20 (3), 327.
- Boto, E., Bowtell, R., Krüger, P., Fromhold, T.M., Morris, P.G., Meyer, S.S., Barnes, G.R., Brookes, M.J., 2016. On the potential of a new generation of magnetometers for MEG: A beamformer simulation study. *PLOS ONE* 11 (8), e0157655. doi:10.1371/journal.pone.0157655.
- Brodbeck, V., Spinelli, L., Lascano, A.M., Wissmeier, M., Vargas, M.-I., Vulliemoz, S., Pollo, C., Schaller, K., Michel, C.M., Seck, M., 2011. Electroencephalographic source imaging: a prospective study of 152 operated epileptic patients. *Brain* 134 (10), 2887–2897.
- Bronstein, M.M., Bruna, J., LeCun, Y., Szlam, A., Vandergheynst, P., 2017. Geometric deep learning: going beyond Euclidean data. *IEEE Signal Process. Mag.* 34 (4), 18–42.
- Brookes, M.J., Boto, E., Rea, M., Shah, V., Osborne, J., Holmes, N., Hill, R.M., Leggett, J., Rhodes, N., Bowtell, R., 2021. Theoretical advantages of a triaxial optically pumped magnetometer magnetoencephalography system. *NeuroImage* 236, 118025.
- Budker, D., Romalis, M., 2007. Optical magnetometry. *Nat. Phys.* 3 (4), 227.
- Chaloner, K., Verdinelli, I., 1995. Bayesian experimental design: a review. *Stat. Sci.* 273–304.
- Chilès, J.-P., Desassis, N., 2018. Fifty years of kriging. In: *Handbook of Mathematical Geosciences*. Springer, pp. 589–612.
- Crane, K., Weischedel, C., Wardetzky, M., 2017. The heat method for distance computation. *Commun. ACM* 60 (11), 90–99. doi:10.1145/3131280.
- Cressie, N., 1993. *Statistics for Spatial Data*. John Wiley & Sons.
- de Munck, J.C., Vijn, P.C.M., Lopes da Silva, F.H., 1992. A random dipole model for spontaneous brain activity. *IEEE Trans. Biomed. Eng.* 39 (8), 791–804. doi:10.1109/10.148387.
- Eldar, Y., Lindenbaum, M., Porat, M., Zeevi, Y.Y., 1997. The farthest point strategy for progressive image sampling. *IEEE Trans. Image Process.* 6 (9), 1305–1315.
- Faley, M., Dammers, J., Maslennikov, Y., Schneiderman, J., Winkler, D., Koshelets, V., Shah, N., Dunin-Borkowski, R., 2017. High-Tc SQUID biomagnetometers. *Supercond. Sci. Technol.* 30 (8), 083001.
- Fang, Q., Boas, D.A., 2009. Tetrahedral mesh generation from volumetric binary and grayscale images. In: 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro. Ieee, pp. 1142–1145.
- Freeman, W.J., Holmes, M.D., Burke, B.C., Vanhatalo, S., 2003. Spatial spectra of scalp EEG and EMG from awake humans. *Clin. Neurophysiol.* 114 (6), 1053–1068.
- Graichen, U., Eichardt, R., Fiedler, P., Strohmeier, D., Zanow, F., Hauelsen, J., 2015. SPHARA—a generalized spatial Fourier analysis for multi-sensor systems with non-uniformly arranged sensors: application to EEG. *PloS one* 10 (4), e0121741.
- Gramfort, A., Luessi, M., Larson, E., Engemann, D.A., Strohmeier, D., Brodbeck, C., Parkkonen, L., Hämäläinen, M.S., 2014. MNE software for processing MEG and EEG data. *NeuroImage* 86, 446–460.
- Grover, P., Venkatesh, P., 2016. An information-theoretic view of EEG sensing. *Proc. IEEE* 105 (2), 367–384.
- Guellmar, D., Hauelsen, J., Reichenbach, J.R., 2010. Influence of anisotropic electrical conductivity in white matter tissue on the EEG/MEG forward and inverse solution. A high-resolution whole head simulation study. *NeuroImage* 51 (1), 145–163.
- Hämäläinen, M., Hari, R., Ilmoniemi, R.J., Knuutila, J., Lounasmaa, O.V., 1993. Magnetoencephalography-theory, instrumentation, and applications to noninvasive studies of the working human brain. *Rev. Mod. Phys.* 65 (2), 413.
- Hamalainen, M.S., Sarvas, J., 1989. Realistic conductivity geometry model of the human head for interpretation of neuromagnetic data. *IEEE Trans. Biomed. Eng.* 36 (2), 165–171.
- Hedrich, T., Pellegrino, G., Kobayashi, E., Lina, J.-M., Grova, C., 2017. Comparison of the spatial resolution of source imaging techniques in high-density EEG and MEG. *NeuroImage* 157, 531–544.
- Hill, R.M., Boto, E., Rea, M., Holmes, N., Leggett, J., Coles, L.A., Papastavrou, M., Everton, S., Hunt, B., Sims, D., et al., 2020. Multi-channel whole-head OPM-MEG: helmet design and a comparison with a conventional system. *NeuroImage* 116995.
- Htet, A.T., Burnham, E.H., Noetscher, G.M., Pham, D.N., Nummenmaa, A., Makarov, S.N., 2019. Collection of CAD human head models for electromagnetic simulations and their applications. *Biomed. Phys. Eng. Express* 5 (6), 067005.
- Iivanainen, J., Stenroos, M., Parkkonen, L., 2017. Measuring MEG closer to the brain: performance of on-scalp sensor arrays. *NeuroImage* 147, 542–553.
- Jacobson, A., et al., 2018. gptoolbox: Geometry Processing Toolbox. <http://github.com/alecjacobson/gptoolbox>.
- Iivanainen, J., Zetter, R., Parkkonen, L., 2020. Potential of on-scalp MEG: Robust detection of human visual gamma-band responses. *Hum. Brain Mapp.* 41 (1), 150–161.
- Jerri, A.J., 1977. The Shannon sampling theorem-its various extensions and applications: a tutorial review. *Proc. IEEE* 65 (11), 1565–1596.
- Kailath, T., Sayed, A.H., Hassibi, B., 2000. *Linear Estimation*. Prentice Hall.
- Kemppainen, P., Ilmoniemi, R., 1989. Channel capacity of multichannel magnetometers. In: *Advances in Biomagnetism*. Springer, pp. 635–638.
- Krause, A., Singh, A., Guestrin, C., 2008. Near-optimal sensor placements in gaussian processes: theory, efficient algorithms and empirical studies. *J. Mach. Learn. Res.* 9 (Feb), 235–284.
- Levy, B., 2006. Laplace-Beltrami eigenfunctions towards an algorithm that “understands” geometry. In: *IEEE International Conference on Shape Modeling and Applications* 2006 (SMI’06). IEEE, pp. 13–13.
- Lew, S., Sliva, D.D., Choe, M.S., Grant, P.E., Okada, Y., Wolters, C.H., Hamalainen, M.S., 2013. Effects of sutures and fontanels on MEG and EEG source analysis in a realistic infant head model. *NeuroImage* 76, 282–293.
- Lindley, D.V., 1956. On a measure of the information provided by an experiment. *Ann. Math. Stat.* 27 (4), 986–1005.
- Loeve, M., 1978. *Probability Theory II*, fourth ed. Springer.
- Malmivuo, J., Plonsey, R., 1995. *Bioelectromagnetism: Principles and Applications of Bioelectric and Biomagnetic Fields*. Oxford University Press, USA.
- McEwen, J.D., Wiaux, Y., 2011. A novel sampling theorem on the sphere. *IEEE Trans. Signal Process.* 59 (12), 5876–5887.
- Mercer, J., 1909. Functions of positive and negative type, and their connection the theory of integral equations. *Philos. Trans. R. Soc. London Ser. A* 209 (441–458), 415–446.
- Mitchell, M.W., Alvarez, S.P., 2020. Colloquium: quantum limits to the energy resolution of magnetic field sensors. *Rev. Mod. Phys.* 92 (2), 021001.
- Neonen, J., Kajola, M., Simola, J., Ahonen, A., 2004. Total information of multichannel MEG sensor arrays. In: *Proceedings of the 14th International Conference on Biomagnetism (Biomag2004)*, pp. 630–631.
- Nunez, P.L., Nunez, M.D., Srinivasan, R., 2019. Multi-scale neural sources of EEG: genuine, equivalent, and representative. A tutorial review. *Brain Topogr.* 32 (2), 193–214.
- Nunez, P.L., Srinivasan, R., et al., 2006. *Electric Fields of the Brain: The Neurophysics of EEG*. Oxford University Press, USA.
- Nurminen, J., Taulu, S., Neonen, J., Helle, L., Simola, J., Ahonen, A., 2013. Improving MEG performance with additional tangential sensors. *IEEE Trans. Biomed. Eng.* 60 (9), 2559–2566.
- Nurminen, J., Taulu, S., Okada, Y., 2010. Improving the performance of the signal space separation method by comprehensive spatial sampling. *Phys. Med. Biol.* 55 (5), 1491.
- Osborne, J., Orton, J., Alem, O., Shah, V., 2018. Fully integrated standalone zero field optically pumped magnetometer for biomagnetism. In: *Steep Dispersion Engineering and Opto-Atomic Precision Metrology XI*, Vol. 10548. International Society for Optics and Photonics, p. 105481G.
- Pesenson, I. Z., 2014. Multiresolution analysis on compact Riemannian manifolds. *arXiv preprint arXiv:1404.5037*.
- Pesenson, I.Z., 2015. Sampling, splines and frames on compact manifolds. *GEM-Int. J. Geomath.* 6 (1), 43–81.
- Petersen, K. B., Pedersen, M. S., 2012. The matrix cookbook <http://www2.compute.dtu.dk/pubdb/pubs/3274-full.html>.
- Petrov, Y., Nador, J., Hughes, C., Tran, S., Yavuzetin, O., Sridhar, S., 2014. Ultra-dense EEG sampling results in two-fold increase of functional brain information. *NeuroImage* 90, 140–145.
- Qiu, A., Bitouk, D., Miller, M.I., 2006. Smooth functional and structural maps on the neocortex via orthonormal bases of the Laplace-Beltrami operator. *IEEE Trans. Med. Imaging* 25 (10), 1296–1306.
- Reuter, M., Biasotti, S., Giorgi, D., Patané, G., Spagnuolo, M., 2009. Discrete Laplace–Beltrami operators for shape analysis and segmentation. *Comput. Graph.* 33 (3), 381–390.
- Riaz, B., Pfeiffer, C., Schneiderman, J.F., 2017. Evaluation of realistic layouts for next generation on-scalp MEG: spatial information density maps. *Sci. Rep.* 7 (1), 6974.
- Robinson, A.K., Venkatesh, P., Boring, M.J., Tarr, M.J., Grover, P., Behrmann, M., 2017. Very high density EEG elucidates spatiotemporal aspects of early visual processing. *Sci. Rep.* 7 (1), 16248.
- Roth, B.J., Sepulveda, N.G., Wikswo Jr, J.P., 1989. Using a magnetometer to image a two-dimensional current distribution. *J. Appl. Phys.* 65 (1), 361–372.
- Sarvas, J., 1987. Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Phys. Med. Biol.* 32 (1), 11.
- Schlömer, T., Heck, D., Deussen, O., 2011. Farthest-point optimized point sets with maximized minimum distance. In: *Proceedings of the ACM SIGGRAPH Symposium on High Performance Graphics*. ACM, pp. 135–142.
- Schneiderman, J.F., 2014. Information content with low-vs. high-Tc SQUID arrays in MEG recordings: the case for high-Tc SQUID-based MEG. *J. Neurosci. Methods* 222, 42–46.
- Sebastiani, P., Wynn, H.P., 2000. Maximum entropy sampling and optimal Bayesian experimental design. *J. R. Stat. Soc. Ser. B* 62 (1), 145–157.
- Shannon, C., 1949. Communication in the presence of noise. *Proc. IRE* 37 (1), 10–21.
- Slutzky, M.W., Jordan, L.R., Krieg, T., Chen, M., Mogul, D.J., Miller, L.E., 2010. Optimal spacing of surface electrode arrays for brain-machine interface applications. *J. Neural Eng.* 7 (2), 026004.

- Solin, A., Särkkä, S., 2020. Hilbert space methods for reduced-rank Gaussian process regression. *Stat. Comput.* 30 (2), 419–446.
- Srinivasan, R., Nunez, P.L., Tucker, D.M., Silberstein, R.B., Cadusch, P.J., 1996. Spatial sampling and filtering of EEG with spline laplacians to estimate cortical potentials. *Brain Topogr.* 8 (4), 355–366.
- Srinivasan, R., Tucker, D.M., Murias, M., 1998. Estimating the spatial Nyquist of the human EEG. *Behav. Res. Methods Instrum. Comput.* 30 (1), 8–19.
- Stark, H., Woods, J.W., 1986. *Probability, Random Processes, and Estimation Theory for Engineers*. Prentice-Hall, Inc..
- Stenroos, M., Hunold, A., Haueisen, J., 2014. Comparison of three-shell and simplified volume conductor models in magnetoencephalography. *NeuroImage* 94, 337–348.
- Stenroos, M., Mäntynen, V., Nenonen, J., 2007. A matlab library for solving quasi-static volume conduction problems using the boundary element method. *Comput. Methods Programs Biomed.* 88 (3), 256–263.
- Stenroos, M., Nummenmaa, A., 2016. Incorporating and compensating cerebrospinal fluid in surface-based forward models of magneto- and electroencephalography. *PLOS ONE* 11 (7), 1–23.
- Stenroos, M., Sarvas, J., 2012. Bioelectromagnetic forward problem: isolated source approach revis (it) ed. *Phys. Med. Biol.* 57 (11), 3517.
- Taulu, S., Kajola, M., 2005. Presentation of electromagnetic multichannel data: the signal space separation method. *J. Appl. Phys.* 97 (12), 124905.
- Tierney, T.M., Mellor, S., O’Neil, G.C., Holmes, N., Boto, E., Roberts, G., ... Barnes, G.R., 2020. Pragmatic spatial sampling for wearable MEG arrays. *Scientific reports* 10 (1), 1–11.
- Vaidyanathan, C., Buckley, K.M., 1997. A sampling theorem for eeg electrode configuration. *IEEE Trans. Biomed. Eng.* 44 (1), 94–97.
- Vorwerk, J., Cho, J.-H., Rampp, S., Hamer, H., Knösche, T.R., Wolters, C.H., 2014. A guideline for head volume conductor modeling in EEG and MEG. *NeuroImage* 100, 590–607.
- Williams, C.K., Rasmussen, C.E., 2006. *Gaussian Processes for Machine Learning*. MIT Press Cambridge, MA.
- Wilson, H., Vrba, J., 2007. Comparison of MEG arrays-revisited. In: *International Congress Series*, Vol. 1300. Elsevier, pp. 619–622.
- Windhoff, M., Opitz, A., Thielscher, A., 2013. Electric field calculations in brain stimulation based on finite elements: an optimized processing pipeline for the generation and usage of accurate individual head models. *Hum. Brain Mapp.* 34 (4), 923–935.
- Wolters, C.H., Anwander, A., Tricoche, X., Weinstein, D., Koch, M.A., MacLeod, R.S., 2006. Influence of tissue conductivity anisotropy on EEG/MEG field and return current computation in a realistic head model: a simulation and visualization study using high-resolution finite element modeling. *NeuroImage* 30 (3), 813–826.
- Xia, H., Ben-Amar Baranga, A., Hoffman, D., Romalis, M., 2006. Magnetoencephalography with an atomic magnetometer. *Appl. Phys. Lett.* 89 (21), 211104.