Automated search of stimulation targets with closed-loop transcranial magnetic stimulation

Aino E. Tervo[1,2], Johanna Metsomaa[1,2,3], Jaakko O. Nieminen[1,2]*, Jukka Sarvas[1,2], Risto J. Ilmoniemi[1,2]

[1]*Department of Neuroscience and Biomedical Engineering, Aalto University School of Science, Espoo, Finland*
[2]*BioMag Laboratory, HUS Medical Imaging Center, University of Helsinki and Helsinki University Hospital, Helsinki, Finland*
[3]*Department of Neurology & Stroke and Hertie Institute for Clinical Brain Research, University of Tübingen, Tübingen, Germany*

**Corresponding author; E-mail: jaakko.nieminen@aalto.fi; Postal address: Department of Neuroscience and Biomedical Engineering, Aalto University School of Science, P.O. Box 12200, FI-00076 AALTO, Finland*

Abstract

Transcranial magnetic stimulation (TMS) protocols often include a manual search of an optimal location and orientation of the coil or peak stimulating electric field to elicit motor responses in a target muscle. This target search is laborious, and the result is user-dependent. Here, we present a closed-loop search method that utilizes automatic electronic adjustment of the stimulation based on the previous responses. The electronic adjustment is achieved by multi-locus TMS, and the adaptive guiding of the stimulation is based on the principles of Bayesian optimization to minimize the number of stimuli (and time) needed in the search. We compared our target-search method with other methods, such as systematic sampling in a predefined cortical grid. Validation experiments on five healthy volunteers and further offline simulations showed that our adaptively guided search method needs only a relatively small number of stimuli to provide outcomes with good accuracy and precision. The automated method enables fast and user-independent optimization of stimulation parameters in research and clinical applications of TMS.

1 Introduction

Neurons in the brain can be excited by transcranial magnetic stimulation (TMS). In TMS, a strong current pulse is fed into a coil placed on the scalp to induce an electric field (E-field) in the cortex (Barker *et al.* 1985). In addition to its use in neuroscience (Lisanby *et al.* 2000, Valero-Cabré *et al.* 2017), TMS is increasingly used in clinical applications ranging from preoperative mapping of motor and speech areas to treatment of various brain disorders such as depression and pain (see, *e.g.*, Lefaucheur and Picht 2016, Lefaucheur *et al.* 2014).

A typical TMS session starts with searching for the optimal stimulation parameters, such as location and orientation of the coil or the E-field maximum, to activate a muscle under investigation most effectively. These optimal stimulation parameters are often referred to as the motor hotspot and defined as the stimulation location and orientation that elicit the largest motor evoked potentials (MEP) measured by electromyography (EMG) (Rossini *et al.* 2015), but the shortest MEP latency or the lowest motor threshold can also be used in the definition (Rossini *et al.* 1994). The motor threshold (MT) is often defined as the stimulation intensity that produces an MEP exceeding a predefined amplitude with a probability of 50%. The E-field at the cortical motor hotspot due to stimulation with MT intensity may serve as a reference when adjusting the stimulation intensity at any cortical site. Therefore, the target search in the motor cortex is often performed even when the actual stimulation site is outside the primary motor cortex.

Optimal stimulation parameters are often searched for by measuring a collection of MEPs when stimulating the cortex around the expected motor representation area, with manually varied stimulation parameters: the location, intensity, and direction of the maximum E-field. The challenge is that, even with fixed stimulation parameters, the MEP amplitude is a random variable, *i.e.*, it varies significantly from stimulation to stimulation due to, for example, excitability fluctuations along the corticospinal tract (Kiers *et al.* 1993). Therefore, target search is laborious and time-consuming. The target search involves also subjective decision making based on the operator's experience, which makes its accuracy and repeatability (over operators) questionable. Sometimes the target may reside in an unexpected location (Ahdab *et al.* 2016, Bulubas *et al.* 2016), in which case the operator's expert opinion may lead to biased localization results.

To make the target search less user-dependent, Meincke *et al.* (2016) and Harquel *et al.* (2017) automated the process with a robotically controlled TMS coupled in a closed loop with EMG feedback. Meincke *et al.* (2016) defined the target as the stimulation coil location leading to the lowest MT; their algorithm started with finding an MEP-positive area within a search grid and continued with evaluating the MT at each MEP-positive site. They reported millimeter-scale repeatability in target localization with an automated protocol that took approximately an hour to complete. Harquel *et al.* (2017) developed an algorithm to find the location maximizing the MEP amplitude using a probabilistic Bayesian approach in which MEP amplitudes were modeled with a 2-dimensional Gaussian function and choosing stimulation sites based on the estimated decrease in entropy. However, both algorithms require a large number of MEPs and are developed to search for the target location with a fixed stimulation orientation (45° from the midsagittal line). The optimal orientation has been shown to differ between hand muscles (Bashir *et al.* 2013) and may also differ by tens of degrees across individuals (Balslev *et*

al. 2007). Therefore, in addition to the stimulation location, the stimulation orientation should be optimized in the target search. Moreover, the performance of both of the above-mentioned algorithms was tested by varying the coil location in a grid with a 7-mm spacing. Such a grid is relatively sparse, and it might be beneficial to carry out the target search in a denser grid.

To overcome these limitations, we developed an algorithm we named BOOST (Bayesian Optimization Of Stimulation Targeting) to automatically find the optimal stimulation parameters, location and orientation, eliciting the largest MEPs. The automated search with BOOST is an iterative process, in which the search result, *i.e.*, the estimate of the optimal target, is updated in each iteration based on the already collected MEPs. Due to the probabilistic nature of the target search, we approach it with Bayesian optimization (see, *e.g.*, Shahriari *et al.* 2016). More specifically, we apply Gaussian process regression (Rasmussen and Williams 2006) to model the MEP responses as a function of the stimulation location and orientation. An advantage of this model is that we make no strong assumptions about the shape of the MEP response function. We present two versions of the BOOST algorithm: (1) in KG-BOOST, the stimulation is adaptively guided with a knowledge-gradient method (Frazier *et al.* 2009, Scott *et al.* 2011, Frazier and Wang 2016), which efficiently optimizes noisy functions (Picheny *et al.* 2013); (2) in Grid-BOOST, MEPs are sampled in a predetermined grid without adaptive guiding. KG-BOOST is the method that we recommend, whereas Grid-BOOST was implemented for comparison.

We evaluated the performance of different versions of the BOOST algorithm in one-dimensional cases with multi-locus TMS (mTMS), which allows electronic adjustment of the stimulation location and orientation without moving the transducer (Koponen *et al.* 2018a, Nieminen *et al.* 2019), making closed-loop stimulation fast end effortless. We hypothesized that, to achieve a given level of performance in the target search, intelligent sampling with knowledge gradient (KG-BOOST) requires less samples than systematic stimulation (Grid-BOOST) in an evenly spaced grid. We also expected KG-BOOST to perform better than Grid-BOOST when the same number of stimuli are used.

2 Methods

In this section, we describe the BOOST algorithm for an automated search for the optimal TMS parameters. In addition, we present the experimental set-up and the measurement protocol used to test the performance of the algorithm.

2.1 Algorithm

First, we present the BOOST algorithm and underlying mathematical models in a general form that allows the search of TMS targets in a multidimensional space. Then, we present a one-dimensional version of the algorithm in detail and the model parameters we used to validate it.

2.1.1 Bayesian optimization with Gaussian processes

Our algorithm is designed to find stimulation parameters that produce the largest motor responses. We model the nth MEP amplitude as

$$y_n = f(\mathbf{x}_n) + \varepsilon_n, \tag{1}$$

where $f(\mathbf{x})$ is an unknown MEP response function that depends on D variables $\mathbf{x} = [x_1, \dots, x_D]^{\mathrm{T}}$. Here, $\mathbf{x}$ can include, for example, the estimated location and orientation of the E-field maximum in the cortex. ε_n represents additive noise and is assumed to follow the normal distribution as $\varepsilon_n \sim \mathcal{N}(0, \lambda_n^2)$, with variance λ_n^2 describing the variability of the MEP responses. Furthermore, ε_n for different n are assumed to be statistically independent of each other and, for simplicity, also of $\mathbf{x}_n$. The measured MEP response, y_n, can be considered to be a random variable, $f(\mathbf{x}_n)$ being the mean or expected value of this noisy sample of f at $\mathbf{x}_n$. Mathematically, the problem of finding the maximum of the MEP response function can be formulated as

$$\mathbf{x}^{\max} = \arg\max_{\mathbf{x}} f(\mathbf{x}), \tag{2}$$

where $\mathbf{x}^{\max}$ contains the stimulation parameters that maximize f. We assume that $f(\mathbf{x})$ is a smooth function that indicates the effect of the E-field distribution on the neuronal pool responsible for the muscle contraction. The non-zero part of f informs us about the motor map of the investigated muscle.

The algorithm is based on modeling f with Gaussian process regression (Rasmussen and Williams 2006) given N noisy MEP responses $\mathbf{y} = [y_1, \dots, y_N]^{\mathrm{T}}$. For stimulation parameters $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_N]$, we assume that the unknown values of f are jointly Gaussian

$$\mathbf{f} \sim \mathcal{N}(\boldsymbol{\mu}_0, \mathbf{K}), \tag{3}$$

where $\mathbf{f} = [f(\mathbf{x}_1), \dots, f(\mathbf{x}_N)]^{\mathrm{T}}$, $\boldsymbol{\mu}_0 = [\mu_0(\mathbf{x}_1), \dots, \mu_0(\mathbf{x}_N)]^{\mathrm{T}}$ is a prior mean vector, and the covariance matrix $\mathbf{K}$ contains information on the *a priori* correlation of $f(\mathbf{x}_n)$ and $f(\mathbf{x}_m)$. In this context, we use a squared exponential covariance kernel defined as

$$\mathbf{K}(n,m) = k(\mathbf{x}_n, \mathbf{x}_m) = a_0 \exp\left(-\sum_{d=1}^{D} a_{1,d} \left|x_{n,d} - x_{m,d}\right|^2\right). \tag{4}$$

The model parameter a_0 determines how much f can vary from the prior mean μ_0, as one can see by setting $m = n$ and observing that $k(\mathbf{x}_n, \mathbf{x}_n) = a_0$ is the variance of $f(\mathbf{x}_n)$. The second set of model parameters $a_{1,d}$ $(d = 1, \dots, D)$ determines the smoothness of f, *i.e.*, how quick changes there can be in each dimension d. The exponential term in Eq. (4) ensures that the correlation of $f(\mathbf{x}_n)$ and $f(\mathbf{x}_m)$ is large when the stimulation parameters $\mathbf{x}_n$ and $\mathbf{x}_m$ (with the elements $x_{n,d}$ and $x_{m,d}$, respectively) are close to each other and $a_{1,d}$ are small. On the other hand, as $a_{1,d}$ get large, $\mathbf{K}$ tends to a diagonal matrix, the samples $f(\mathbf{x}_n)$ becoming mutually less correlated.

With the help of the described model and Bayesian inference, one can formulate the posterior probability distribution of $f(\mathbf{x})$ at any point $\mathbf{x}$ given the measured N responses in vector $\mathbf{y}$ and the corresponding stimulation parameters $\mathbf{X}$:

$$f(\mathbf{x}) \mid \mathbf{y}, \mathbf{X} \sim \mathcal{N}(\mu_N(\mathbf{x}), \sigma_N^2(\mathbf{x})). \tag{5}$$

In this formula, the so-called posterior mean $\mu_N(\mathbf{x})$ can be calculated as

$$\mu_N(\mathbf{x}) = \mu_0(\mathbf{x}) + \mathbf{k}(\mathbf{x}, \mathbf{X})(\mathbf{K} + \boldsymbol{\Lambda})^{-1}(\mathbf{y} - \boldsymbol{\mu}_0), \tag{6}$$

where $\mathbf{k}(\mathbf{x}, \mathbf{X}) = [k(\mathbf{x}, \mathbf{x}_1), \dots, k(\mathbf{x}, \mathbf{x}_N)]$ contains the covariances between $\mathbf{x}$ and $\mathbf{x}_1, \dots, \mathbf{x}_N$ similarly to Eq. (4), and $\boldsymbol{\Lambda}$ is a diagonal matrix with $\lambda_1^2, \dots, \lambda_N^2$ on its diagonal. The location of the posterior mean maximum $\mu_N^{\max} = \max(\mu_N(\mathbf{x}))$ provides an estimate of the

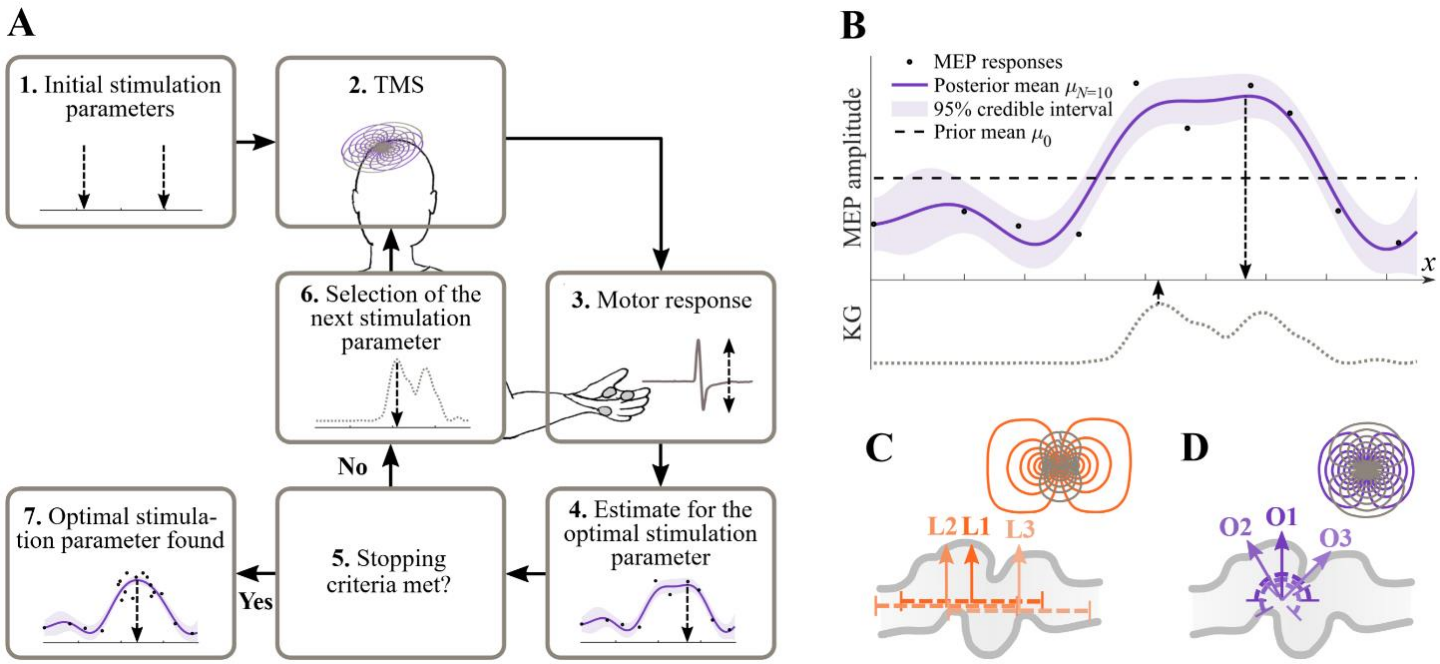

Figure 1. Bayesian optimization of stimulation targeting. A: Algorithm flowchart. B: Example of the estimation of the optimal stimulation target after 10 MEP responses (black dots). The purple solid line shows the posterior mean curve, and the shaded area depicts its 95% credible interval. The current estimate of the optimal stimulation target is where this curve reaches its maximum (downward-facing arrow). The black dashed line, being the mean of the posterior mean curve, depicts the prior mean for the next iteration. The next stimulus in KG-BOOST would be given with the parameter that maximizes the knowledge-gradient curve (gray dotted line; upward-facing arrow). C: Three placements (L1–L3) of the translation transducer over the primary motor cortex for 1D location search. The coil windings of the transducer are visualized in the top-right corner. D: The transducer placements for the orientation search (O1–O3) and the coil windings of the rotation transducer. In C and D, the dashed lines show the search space and the arrows indicate the location and orientation of the E-field maximum in the reference origin for each transducer placement. In each case, the stimulation of the reference origin is realized with the lower coil (orange or purple), and the stimulation of the other targets is achieved by feeding suitable current combinations to both of the overlapping coils.

optimal stimulation parameters $\mathbf{x}^{\max}$. One can also estimate the variance of $\sigma_N^2(\mathbf{x})$ in Eq. (5):

$$\sigma_N^2(\mathbf{x}) = k(\mathbf{x},\mathbf{x}) - \mathbf{k}(\mathbf{x},\mathbf{X})(\mathbf{K}+\mathbf{\Lambda})^{-1}\mathbf{k}(\mathbf{x},\mathbf{X})^{\mathrm{T}}. \qquad (7)$$

The sampling in the optimization process can be performed systematically in a grid of stimulation parameters as we do in the Grid-BOOST version of the algorithm. Another option is to utilize so-called acquisition functions that provide suggestions for the next sampling parameters $\mathbf{x}_{N+1}$ and help finding the optimum with a smaller number of samples than with the pre-determined grid approach. For this purpose, we apply the knowledge gradient (KG) that guides KG-BOOST optimization and is computed as

$$\mathrm{KG}(\mathbf{x}_{N+1}) = E[\mu_{N+1}^{\max}(\mathbf{x}_{N+1})] - \mu_N^{\max}\,, \qquad (8)$$

where $\mu_{N+1}^{\max}(\mathbf{x}_{N+1})$ represents the maximum of the posterior mean function if we get one extra sample y_{N+1} corresponding to the parameters $\mathbf{x}_{N+1}$. In Eq. (8), $\mu_{N+1}^{\max}(\mathbf{x}_{N+1})$ is a random variable depending on the random variable y_{N+1} and E stands for the mean operator. $\mathrm{KG}(\mathbf{x}_{N+1})$ thus tells us how much the maximum of the posterior mean is expected to change if a new sample were collected with parameters $\mathbf{x}_{N+1}$. The optimally chosen next stimulation parameters lie where $\mathrm{KG}(\mathbf{x}_{N+1})$ reaches its maximum. More details about the knowledge-gradient computation can be found in the Appendix. Note that there are also other methods to guide adaptive sampling, but we chose the knowledge-gradient method due to its reported good performance in optimizing noisy functions (Picheny *et al.* 2013).

2.1.2 Algorithm for finding an optimal stimulation target in one dimension

The algorithm presented above works in multiple dimensions; here, we describe its one-dimensional version to find automatically either the optimal location of the peak E-field on a line segment or the optimal E-field orientation. Figure 1A depicts the flowchart of the algorithm. Several choices for the algorithm, such as the method for choosing the first stimulation parameters, were made intuitively. Most of these choices are justified by the fact that they resulted in a well-working algorithm. The automated search begins with determining the stimulation parameters for two initial samples within the search space $[-L/2, L/2]$ that covers the selected range of stimulation locations or orientations (Step 1). The first initial sample is randomly chosen within the first half of the search space, *i.e.*, $x_1 \in [-L/2, 0]$. The second one is taken at location $x_2 = x_1 + L/2$. Next, a TMS pulse (or two pulses directly after Step 1) is given according to the determined stimulation parameter (Step 2), and the resulting MEP is measured from the target muscle (Step 3). For our analysis, we define the MEP amplitude y_n as the base-10 logarithm of the measured peak-to-peak amplitude (in microvolts), to meet better the Gaussian assumption of the MEP variability and to suppress large outlier MEPs, which could otherwise lead to an erroneous final result.

In Step 4, we compute the estimate for the posterior mean with Eq. (6). Since the MEP variability differs between individuals, we adjust most of the model parameters adaptively based on the data gathered in the optimization process. As a prior mean for the first iteration, *i.e.*, after the two initial samples, we use a constant function, the output value of which is the average of the two initial responses. After the first iteration, the prior mean is still constant but now with an output value that is the average of the current posterior mean curve: $\mu_0(\mathbf{x}) = \overline{\mu_N(\mathbf{x})}$, where $\overline{\mu_N(\mathbf{x})}$ denotes the mean of the posterior mean curve (see the black dashed line in Fig. 1B). In the first iteration, we set the variability parameter as $a_0 = \max_{n=1,2}(|y_n - \bar{y}|^2)$, where $\bar{y}$ denotes the mean of the gathered MEPs. If the two initial responses happen to have the same value, we set $a_0 = (\log_{10}5)^2 = 0.49$. In the subsequent iterations, we define $a_0 = \max(|\max(\mathrm{CI}_{95\%,\mathrm{u}}(\mathbf{x})) - \mu_0(\mathbf{x})|^2, |\mu_0(\mathbf{x}) - \min(\mathrm{CI}_{95\%,\mathrm{l}}(\mathbf{x}))|^2)$, where $\mathrm{CI}_{95\%,\mathrm{u}}(\mathbf{x}) = \mu_N(\mathbf{x}) + 1.96\sqrt{\sigma_N^2(\mathbf{x})}$ and $\mathrm{CI}_{95\%,\mathrm{l}}(\mathbf{x}) = \mu_N(\mathbf{x}) - 1.96\sqrt{\sigma_N^2(\mathbf{x})}$ are the upper and lower limits of the 95% Bayesian credible interval of the posterior mean estimate, respectively (see an example of the credible interval in Fig. 1B). We define the smoothness parameter as $a_1 = k^2\pi^2/(2L^2)$, where k tells how many times f is expected to cross its mean value within the search space (see Rasmussen and Williams 2006, page 81). Here, we assume that $k = 2$. The MEP variance λ_n^2 is chosen to be constant everywhere. In the first iteration, $\lambda_n^2 = \lambda^2 = 1/2\sum_{n=1}^{2}|y_n - \bar{y}|^2$, which is the variance of the elements in $\mathbf{y}$. In the following iterations, $\lambda_n^2 = \lambda^2 = 1/N\sum_{n=1}^{N}|y_n - \mu_N(\mathbf{x}_n)|^2$, *i.e.*, the variance of the differences between the measured MEPs and the posterior mean curve. In each iteration, the estimate of the optimal stimulation target is where the posterior mean reaches its maximum.

In Step 5, the algorithm checks the stopping criteria. If they are not met, the next sampling point is determined (Step 6). The KG-BOOST algorithm stops if at least 14 responses have been collected and if, during the past eight iterations, the estimate of the optimal stimulation target has changed no more than 2 mm or 12° for the location and orientation search, respectively, or when 30 responses have been collected. These stopping criteria were formed based on preliminary convergence evaluations on test data. If the stopping criteria are not met, the next sampling point is chosen by evaluating the knowledge-gradient function (Eq. (8)) and by finding its maximum. If the knowledge gradient has the same value for all sampling points, we randomly pick the next stimulation parameter from 2–5 mm or 12–30° distance from the current estimate of the optimal stimulation target. We repeat Steps 2–6 until the stopping criteria are met. In Grid-BOOST, we take one sample in each point of an equally spaced grid in random order with no adaptive stopping criteria, sampling until the whole search space is covered systematically.

2.2 Data acquisition

Five healthy subjects volunteered for the study (aged 26–35 years, two males). All subjects were right-handed according to the Edinburgh inventory (Oldfield 1971). Prior to the measurements, each subject signed an informed consent. The study was approved by the Coordinating Ethics Committee of the Helsinki University Hospital and was carried out in accordance with the Declaration of Helsinki.

TMS was administered with two different transducers connected to our in-house-developed mTMS system (Koponen *et al.* 2018a). One of the transducers comprises a figure-of-eight coil and an overlapping oval coil (Fig. 1C; Koponen *et al.* 2018a). With this translation transducer, we could electronically shift the location of the calculated E-field maximum along a 30-mm-long line segment in the cortex. The E-field in the cortex was calculated using a spherical head model with an 85-mm radius, the cortex assumed to be at 15-mm depth from the head surface. We had 31 possible locations (symmetrically around the reference origin with 1-mm spacing) of the E-field maximum along this line segment. The other transducer, with two overlapping figure-of-eight coils (Fig. 1D), allows electronic adjustment of the orientation of the maximum E-field (de Oliveira e Souza 2018). We restricted the possible stimulation orientations in the spherical head model to be within a 180° interval centered around the reference origin, with neighboring orientations separated by 1°. Here, the reference origin (0 mm/0°) means the location or orientation of the maximum E-field resulting from the stimulation with only the lower of the two overlapping coils (Figs. 1C, D). Thus, the reference origin moves together with the transducer. The applied pulse waveforms were monophasic (60-µs rise time, 30-µs hold period; Koponen et al. 2018b) and the interstimulus interval was randomized between 4 and 6 s.

The position of the mTMS transducers and the head of the subject were tracked with a neuronavigation system (eXimia 3.2, Nexstim Plc, Finland). For image-based guiding, we had T1-weighted magnetic resonance images of each subject. When needed, the position of the transducer with respect to the subject's head was kept fixed with the help of the neuronavigation system, which allowed stimulation only when the transducer location was within 2 mm and all rotation angles less than 2° from their target values.

The motor responses were measured with surface EMG integrated in the eXimia 3.2 system (500-Hz low-pass filtering, 3-kHz sampling frequency). The silver/silver-chloride surface electrodes (Ambu Neuroline 720, Ambu A/S, Denmark) were in a bipolar arrangement with the active electrode placed over the muscle belly of the right *first dorsal interosseus* (FDI) and the reference electrode on the second proximal phalange. The ground electrode was placed on the back of the hand. The eXimia system analyzed the evoked responses and displayed the peak-to-peak amplitudes and latencies of the MEPs in real time. To get the MEP data into our algorithms programmed with Matlab (The MathWorks, Inc., USA), we imported the video stream of the eXimia system to our control computer in real time with a USB video grabber (DVI2USB 3.0, Epiphan Systems Inc., Canada). From the video stream, we extracted the MEP-amplitude and MEP-latency values as reported by the eXimia system. We also analyzed the baseline EMG signal 5–200 ms before the TMS pulse for real-time rejection of responses with muscle preactivation. We accepted an MEP if its onset latency was 15–30 ms and if the baseline EMG signal was within ±10 µV (when determining the MT) or ±15 µV (when running the KG-BOOST or Grid-BOOST algorithms). If these conditions were not met, we repeated the stimulation with the same parameters until the MEP response was acceptable. Since the automatic MEP analyzer of the eXimia system sometimes missed small responses, giving just 0 µV as their amplitude, we replaced for data analysis each 0-µV MEP amplitude with a random amplitude drawn from a uniform distribution with an interval of 5–15 µV.

Each subject had two measurement sessions, conducted on different days. At the beginning of the first session, we manually located the optimal stimulation target of the FDI muscle with the figure-of-eight coil of the translation transducer. For this, we delivered TMS pulses to the left primary motor cortex, varying the target location with millimeter-level steps around the hand-knob area. The stimulation intensity was first about 70 V/m, then adjusted so that the maximal MEP amplitude would be approximately 1 mV. The ISI was about 5 s. After delivering several tens of pulses around the hand-knob region to outline the MEP-positive area, we visually evaluated the distribution of the MEP responses and selected one target approximately from the center of the area showing the largest MEPs. During the manual search, the estimated orientation of the peak E-field was kept perpendicular to the overall orientation of the precentral gyrus. Next, we determined the resting MT (rMT) for the FDI muscle with a maximum-likelihood method (Awiszus 2003), applying 20 pulses with different intensities while having 50 µV as a threshold for MEP acceptance. For the rest of the two sessions, the stimulation intensity was set to 110% rMT.

In the first session, we performed the automated search for the optimal stimulation location with the KG-BOOST and Grid-BOOST algorithms described in Section 2.1.2. We positioned the transducer in three different locations to test whether the algorithm can locate the optimal stimulation site regardless of its placement within the search space. The first placement corresponded to the manually found FDI target (placement L1), the second was ~5 mm to the medial (placement L2) and the third ~8 mm to the lateral direction (placement L3) from L1 (see Fig. 1C). The transducer was kept fixed at L1–L3 and the peak E-field was electronically adjusted to one of the 31 possible locations according to the algorithm. We repeated both KG-BOOST and Grid-BOOST seven times for each transducer placement (L1–L3), resulting in 21 repetitions per subject. The transducer placements and the utilized version of the algorithm were applied in a pseudorandom order. The experimenters were aware of the transducer placement as well as the applied algorithm version during the experiments. With both KG-BOOST and Grid-BOOST, the estimated posterior mean curve computed with Gaussian process regression (Eq. (6)) was calculated on the 30-mm line segment with a grid spacing of 0.25 mm.

In the second session, we conducted the automated search for the optimal stimulation orientation. We set three transducer placements as follows: the placement O1 corresponded to L1, the placement O2 was oriented ~30° counterclockwise, and the placement O3 ~45° clockwise with respect to O1 (see Fig. 1D). The sampling in Grid-BOOST consisted of 31 pulses with 6° steps ranging from −90° to 90° around the reference origin. In KG-BOOST, we had the same 180°-wide search space with the possible stimulation orientations separated by 1° steps. We repeated both KG-BOOST and Grid-BOOST seven times with transducer placements O1–O3, resulting in 21 repetitions for each subject. The posterior mean curve was computed with a grid spacing of 0.5°.

2.3 Data analysis

In this Section, we first explain how we simulated other search methods using the measured data. Then, we show how we evaluated the accuracy of different search methods by comparing the optimization results with the ground truth and how we determined the precision as the deviation in the search outcomes. In addition, we present details of the statistical testing comparing the performance of KG-BOOST with the other methods. The investigators were not blinded when analyzing the data.

2.3.1 Other search methods

To complement the results obtained directly from our experiments, we simulated the performance of three other sampling methods using the MEPs collected in the KG-BOOST and Grid-BOOST searches, sampling from these data without replacement.

First, we conducted sparser grid sampling (sparse Grid-BOOST) with the data collected in the original Grid-BOOST searches (that we call dense Grid-BOOST from now on). In sparse Grid-BOOST, the number of samples collected was equal to that of the corresponding KG-BOOST search repetition (14–30 samples per search). We selected every other sample from the denser grid and sampled from this subset of the data in random order until the total number of samples was the same as in the KG-BOOST search. If needed, we took extra samples from the unused half of the data. Since the two data subsets could be constructed in two ways, with the first one including either even or odd indices, we randomized the order in which the two subsets were used. The posterior mean curve corresponding to the sparse Grid-BOOST sampling was computed in the same way as with KG-BOOST and dense Grid-BOOST, *i.e.*, with the parameters presented in Section 2.1.2.

To mimic the target search without any modeling of the MEP responses, we also evaluated a search strategy in which the estimate of the optimal stimulation target coincided with the location of the maximum MEP response. For this, we used the data sampled in the sparse Grid-BOOST searches. We refer to this search strategy as the maximum-MEP method.

For comparison, we simulated the previously reported AutoHS method by Harquel *et al.* (2017) to find the optimal stimulation site. We made three adjustments to the AutoHS method due to differences in the MEP-sampling schemes: (1) Our sampling grid spacing was 1-mm (cortical grid) as opposed to 7 mm (grid of coil locations on the scalp) used in the original study. (2) Our search space was one-dimensional instead of two-dimensional. (3) We allowed sampling at each stimulation site at most once, whereas in Harquel *et al.* (2017) the same location was targeted at most twice. Because AutoHS gathers five MEPs at each iteration at the selected stimulation target and because for some targets we had collected only seven MEPs, we run the method only once for each subject and transducer placement. When searching for the optimal stimulation location with AutoHS, the possible values for the maximum MEP amplitude, the Gaussian width, and the center point of the Gaussian were {100 µV, 300 µV, …, 3900 µV}, {2 mm, 4 mm, ..., 20 mm}, and {−15 mm, −14.75 mm, …, 15 mm}, respectively. For defining the next stimulation target in each iteration, the grid spacing was 1 mm. The first stimulation target and the center point of the prior distribution were always set in the middle of the search space (the reference origin).

We also performed the search for the optimal stimulation orientation with AutoHS although such an application was not described in the original article. Indeed, the shape of the Gaussian function could be expected to model the MEP distribution as a function of the stimulation orientation, too. Here, the possible values for the maximum MEP amplitude, the Gaussian width, and

the center point of the Gaussian were {100 μV, 300 μV, …, 3900 μV}, {12°, 24°, …, 120°}, and {−90°, −89.5°, …, 90°}, respectively. The spacing of the sampling grid was 6°. The prior for the optimal angle was centered around the reference origin as in the location search. The width of the prior distribution was 30°.

2.3.2 Estimation of accuracy and precision

To estimate the bias in the results (*i.e.*, the accuracy) obtained with different search strategies, we first defined the best estimate for the optimal stimulation target (hereafter, the ground truth) for each transducer placement and subject. This was done by first pooling the data from the seven repetitions of the KG-BOOST and Grid-BOOST searches. Each data pool included at least 315 MEP responses from the same spatial/angular distribution. For each of the 30 cases (five subjects, two transducers, three transducer placements), we computed a median curve in a grid with a 1-mm/1° spacing using a sliding window that took into account the responses that were closer than 5 mm (location search) or 30° (orientation search) from the computation point. We defined the ground truth as the location of the maximum of the median curve. If several points of this curve had the same maximum value, we defined the ground truth as their mean location. When calculating the ground truth, we replaced the amplitudes of those MEPs that the eXimia system had originally identified as 0-μV MEPs with the peak-to-peak amplitude of the EMG signal in the time interval of 15–45 ms after the TMS pulse.

We computed the average location/orientation of the search results over the seven repetitions, separately for each search strategy (KG-BOOST, dense and sparse Grid-BOOST, the maximum-MEP method), and compared it with the corresponding ground truth. To determine the group-level accuracy, we computed the mean of these differences in 15 cases (3 transducer placements × 5 subjects). This accuracy measure tells us how close the average result was to the ground truth. With AutoHS, we had in each case only one simulated search result and used its difference from the ground truth when computing the mean accuracy.

To assess the precision (degree of scatter) of each search method, we computed the standard deviation of the corresponding seven final search results and averaged these standard deviations over the transducer placements and subjects. This precision measure describes the repeatability of the outcome of each search method.

2.3.3 Statistical analysis

The difference between the precision/accuracy of KG-BOOST and the other search methods with the same number of search repetitions (*i.e.*, dense and sparse Grid-BOOST, the maximum-MEP method) was tested by permutation statistics as follows. Altogether, 30 test values (5 subjects × 3 transducer placements × 2 search methods under comparison) were randomly divided into two groups 1,000,000 times. The accuracy and precision for both the location and the orientation search were treated separately. For each permutation, we computed the mean value for both groups. We obtained a two-tailed *p*-value as the proportion of permutations for which the absolute value of the difference of means of the permuted groups exceeded the absolute value of the corresponding original difference of means between the sampling methods. The level of statistical significance was set at 0.01. After Bonferroni correction for multiple comparisons (12 comparisons), the corrected significance level was 0.00083.

3 Results

Results of the location and orientation searches are visualized in Fig. 2. Figure 2A shows how the search results for a representative subject with the three different transducer placements L1–L3 are distributed with respect to the ground truth. The three smaller graphs on the right side of Fig. 2A depict an example of a single run with each of the four search methods (KG-BOOST, dense and sparse Grid-BOOST, the maximum-MEP method). Figure 2B shows similar example results for the orientation search. Figures 2C and 2D illustrate the error distributions of the search outcomes, *i.e.*, how far the optimized parameters are from the ground truth, for each subject.

The convergence of the adaptively guided KG-BOOST in the search for the optimal location and orientation is presented in Figs. 3A and 3B, respectively. Figure 3C depicts the performance metrics of the location search over subjects. The mean accuracy (*i.e.*, how far the average search result is from the ground truth) of KG-BOOST was 1.4 mm (range: 0.04–5.1 mm) which is similar to the mean accuracy of 1.5 mm (range: 0.2–5.0 mm) of dense Grid-BOOST ($p = 0.87$). The accuracies of sparse Grid-BOOST (mean: 2.0 mm; range: 0.04–4.4 mm) and the maximum-MEP method (mean: 2.1 mm; range: 0.3–5.3 mm) were slightly worse, although the differences were not statistically significant when compared to the accuracy of KG-BOOST ($p = 0.32$ for the sparse Grid-BOOST and $p = 0.21$ for the maximum-MEP method). The precision, expressing how repeatable the results are, was best with dense Grid-BOOST (mean: 2.7 mm; range: 1.2–5.2 mm). The mean precision of KG-BOOST (mean 3.2 mm; range 1.6–5.3 mm), the maximum-MEP method (mean: 3.4 mm; range: 1.9–6.6 mm) and sparse Grid-BOOST (mean: 3.4 mm; range: 1.7–6.2 mm) were comparable to each other. When comparing KG-BOOST with dense or sparse Grid-BOOST or the maximum-MEP method, the *p*-values were 0.23, 0.68, and 0.72, respectively. With KG-BOOST, sparse Grid-BOOST and the maximum-MEP method, the average number of samples collected was 18 (range: 14–30), which means that on average the search took 1.5 min. With dense Grid-BOOST, we always gathered 31 samples per search, corresponding to an average time of 2.6 min. Moreover, the manually found optimal stimulation location, which corresponds to the reference origin with the transducer position L1, varied on average 3.5 mm (range: 0–6 mm) from the corresponding ground truth. The average number of stimuli in the manual search was 59 (range: 37–78), administered on average in 11 minutes (range: 4–18 min).

The performance metrics of the orientation search can be found in Fig. 3D. The mean accuracy of KG-BOOST (mean: 5.4°; range 1.4–13.6°) was close to the ones of dense Grid-BOOST (mean: 5.1°; range: 0.6–10.3°; $p = 0.82$), sparse Grid-BOOST (mean: 5.8°; range: 0.9–18.8°; $p = 0.78$) and the maximum-MEP method (mean: 7.5°; range; 0–17.1°; $p = 0.21$). The mean precisions of KG-BOOST (mean: 9.7°; range 5.5–15.4°), dense Grid-BOOST (mean: 11.3°; range: 4.2–19.8°), sparse Grid-BOOST (mean: 19.1°; range: 8.3–48.8°) and the maximum-MEP method (mean: 20.2°; range: 10.4–32.9°) varied from each other. We found statistically significant differences in precision between KG-BOOST and both sparse Grid-BOOST ($p = 0.000031$) and the maximum-MEP method ($p = 0.000001$). The difference in precision between KG-BOOST and dense Grid-BOOST was not statistically significant ($p = 0.25$). The average number of samples acquired in the orientation search was 16 (range: 14–30; average

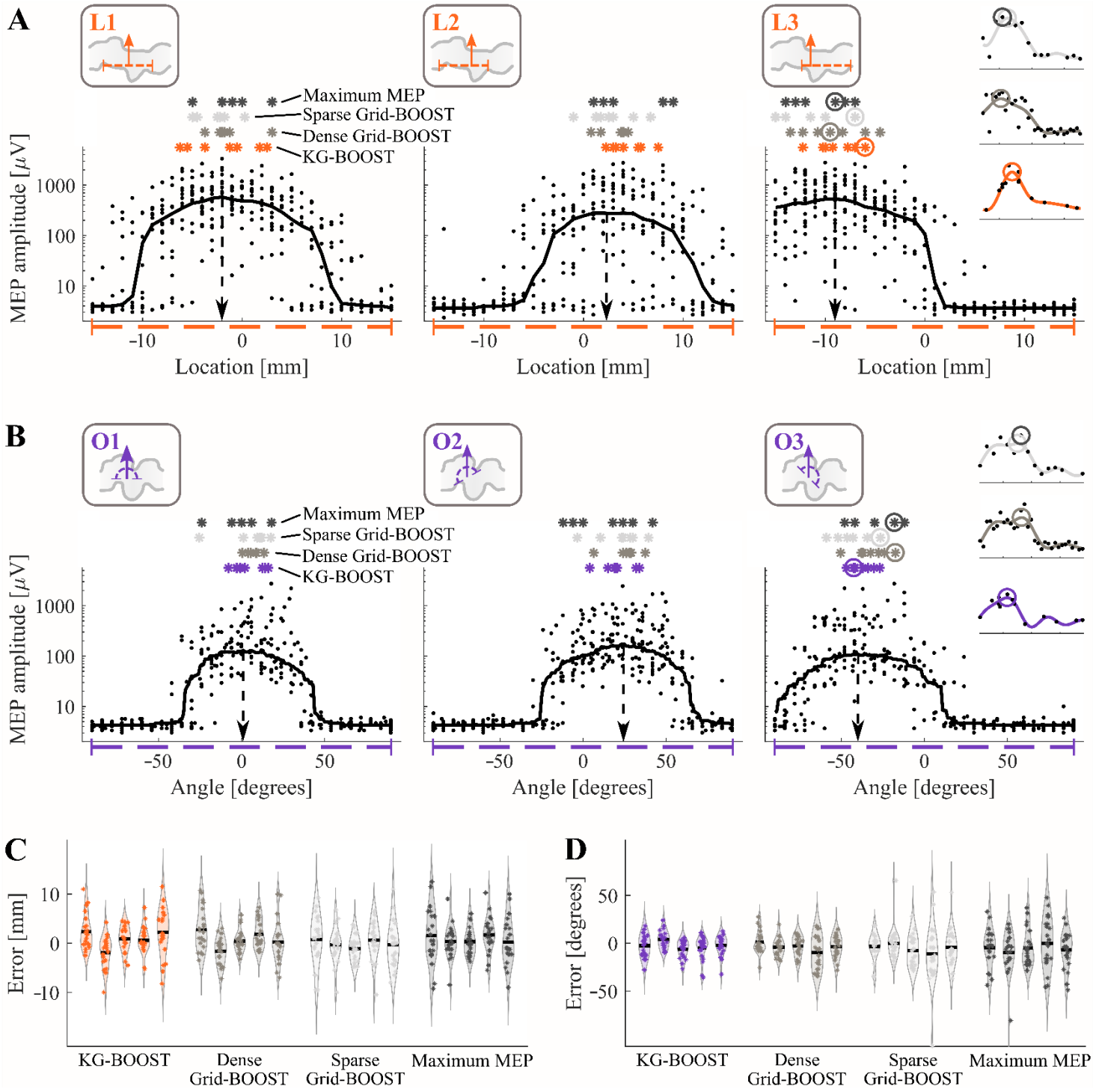

Figure 2. Results of the location and orientation searches. A: Location search with Subject 3 and transducer placements L1–L3. B: Orientation search with Subject 1 and transducer placements O1–O3. The transducer placements with respect to the hand-knob area in the pre-central gyrus are shown in the boxes on the top (the arrows indicate the location/orientation of the manually found target), with the corresponding results panels below them. The black dots show the MEP responses measured during all KG-BOOST and dense Grid-BOOST searches. The black line shows the moving median of these responses, and the location of its maximum (black dashed arrow) is the ground truth. The asterisks show the estimates of the optimal stimulation locations/orientations of the seven repetitions with KG-BOOST (orange/purple), dense (gray) and sparse (light gray) Grid-BOOST and the maximum-MEP method (dark gray), respectively. The three smaller graphs on the right visualize single runs of different search methods with L3/O3. In these graphs, the black dots illustrate the MEP responses, the solid lines depict the posterior mean curves, and the circles indicate the estimated optimal stimulation locations/orientations (the corresponding asterisks in the L3/O3 results panel are circled). C–D: Violin plots visualizing the error (difference from the ground truth) in the search of the optimal location (C) and orientation (D) for each subject, with the data of all transducer placements combined. The asterisks show the error of each search run, and the black lines indicate the mean errors.

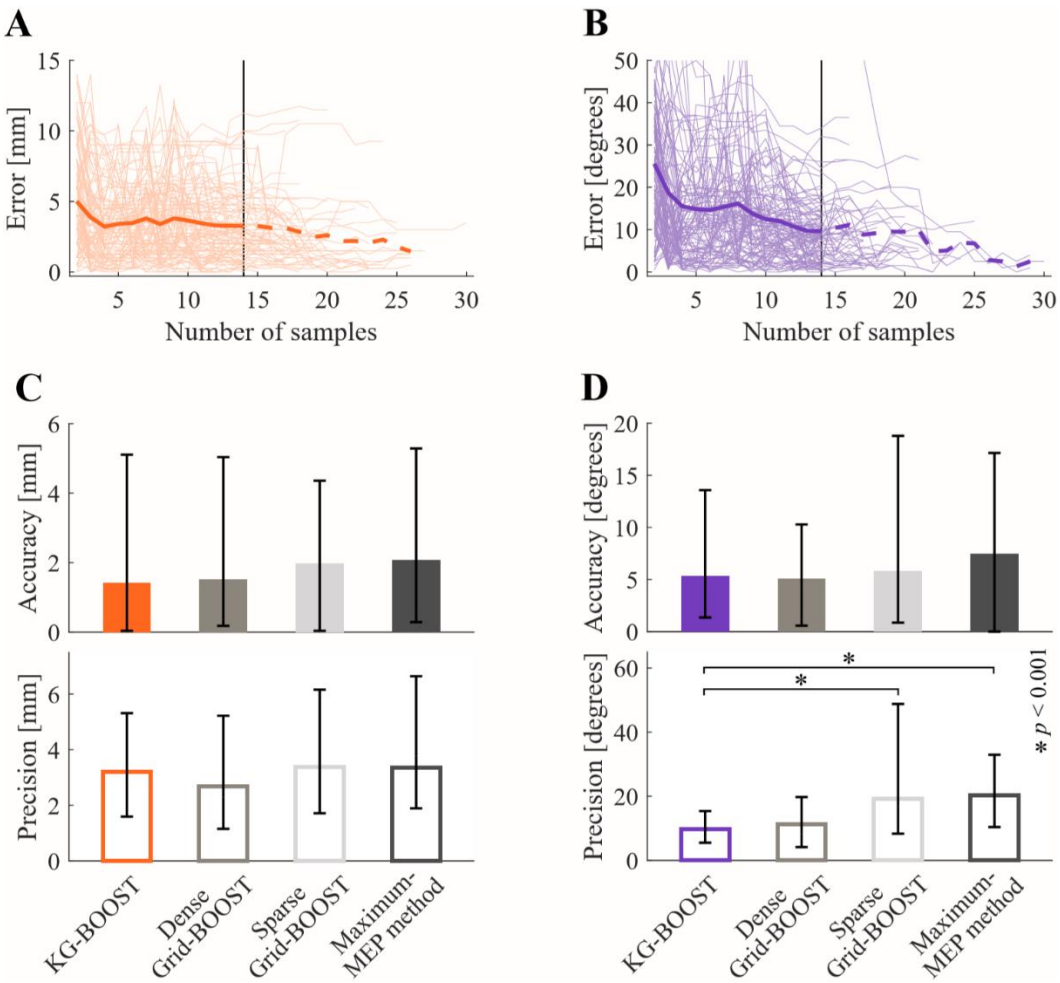

Figure 3. The convergence of KG-BOOST and the performance of the different search methods. A–B: The convergence of KG-BOOST in the search for the optimal stimulation location (A) and orientation (B). The light curves show how far from the ground truth the single runs are as a function of the number of samples. The solid dark curve depicts the average error from the ground truth until the minimum number of samples (14, black vertical line) has been reached. After 14 samples, as many of the runs have already finished, the average error curve (dashed line) includes only the remaining runs (until three or more left). C–D: Accuracy and precision of the different search methods in the search for the optimal stimulation location (C) and orientation (D). The bars depict the mean accuracy/precision of the search results with the black whiskers showing the minimum and maximum accuracy/precision over subjects and transducer placements. In D, the asterisks indicate the statistically significant differences in the precision of sparse Grid-BOOST and the maximum-MEP method compared to the precision of KG-BOOST.

time: 1.3 min) for KG-BOOST, sparse Grid-BOOST, and the maximum-MEP method. With dense Grid-BOOST, the number of samples was always 31 (average time: 2.6 min). In addition, the reference origin of the transducer placement O1, which was set to be perpendicular to the global orientation of the precentral gyrus, varied on average 6.7° (range: 1–13°) from the ground truth.

Our version of the AutoHS algorithm yielded an average error of 1.5 mm (range: 0.04–4.6 mm) and 19.7° (range: 2.1–37.9°) with respect to the ground truth for the location and orientation search, respectively. These values are not directly comparable with the accuracy values presented in Figs. 3C and 3D, since only one iterative search could be simulated for each subject and coil placement with the data available. Therefore, the results of AutoHS are excluded from Figs. 2 and 3 and the statistical analysis. With the AutoHS algorithm, the number of responses required for convergence was on average 53 (range: 20–110) and 62 (range: 20–163) for the location and orientation search, respectively, being about three to four times the number of stimuli used by KG-BOOST.

4 Discussion

We demonstrated that multi-locus TMS and Bayesian optimization can be successfully combined into an automated search of TMS targets. In this context, mTMS enables adjusting the stimulation location and orientation in a closed-loop setting without the need to move a coil, which significantly reduces the laboriousness of TMS. Bayesian optimization provides means to model and guide the stimulation in an effective and user-independent manner.

4.1 Performance of the automated target search

The automated online searches and further offline comparisons revealed that the mean accuracy in the location search was almost the same with all three versions of the BOOST algorithm (KG-BOOST guided with knowledge gradient and Grid-BOOST with dense and sparse sampling grids), when choosing the optimal target based on the maximal MEP response only, and with the

AutoHS method (Harquel *et al.* 2017) (accuracies in the range of 1.4–2.1 mm). In the orientation search, the accuracy of the AutoHS method (19.7°) was worse than that of the other methods (5.1–7.5°). The small accuracy values indicate that the search results were centered almost symmetrically around the ground truth, which is expected behavior for any sensible search method.

The average precision of the location search was similar (2.7–3.4 mm) among the four search strategies for which we were able to compute the precision (for AutoHS, we did not have enough samples for retrieving several independent search results). Thus, there were no statistically significant differences in precision in the location search between KG-BOOST and the other search methods. This can be explained by the fact that the motor maps extended over a large portion of the 30-mm-long search space and the median curves were relatively flat around their maximum (Fig. 2A). Instead, we found differences in the average precision of the optimal orientation search. KG-BOOST and dense Grid-BOOST had similar precisions (9.7° and 11.3°, respectively), while the precision of sparse Grid-BOOST and the maximum-MEP method were significantly worse (19.1° and 20.2°, respectively). These differences in the degree of scatter are not surprising, since the MEP-positive part of the median curve in the search space appeared narrower in the orientation search compared to the corresponding part in the location search. In the orientation search, KG-BOOST and dense Grid-BOOST got enough samples from the maximum area and the slopes of the mean MEP curve whereas sparse Grid-BOOST and the maximum-MEP method got fewer responses around the maximum, leading to larger deviation and, thus, worse precision in the search results.

Considering the efficiency of different search strategies, dense Grid-BOOST always used 31 samples, and AutoHS needed on average 53 and 62 samples, whereas the other methods used 18 and 16 samples on average in the location and orientation search, respectively. Sampling with KG-BOOST resulted in accuracy and precision similar to those of dense Grid-BOOST, but with approximately half of the number of samples. Also, the accuracies were similar (location search) or worse (orientation search) with AutoHS than with KG-BOOST, and AutoHS needed on average three to four times more samples for convergence. These findings indicate that KG-BOOST was more efficient than AutoHS and dense Grid-BOOST, and that sampling in a dense evenly spaced grid wastes samples especially in areas that produce no MEP responses. Results of the orientation search show that adaptive sampling with KG-BOOST led to better precision than placing the same number of samples evenly in the search space. Based on these results, we suggest using intelligent sampling, such as sampling with knowledge gradient (Frazier *et al.* 2009) that we used in KG-BOOST. This would allow one to avoid gathering too much data in areas with no MEPs while sampling adaptively around the maximum of the MEP curve to efficiently get enough information about the optimal target. We anticipate that even larger differences between the sampling methods are expected with larger or higher-dimensional search spaces (such as a two-dimensional location grid with additional variation in orientation), when non-guided grid sampling becomes very time-consuming.

Although the accuracy and precision values are generally good for almost all the search strategies, a single search outcome may still be several millimeters or degrees off from the ground truth regardless of the search method (see examples in Fig. 2). This is mainly due to the unavoidable high variability of the MEP responses that is present also in manual searches. Setting more strict stopping criteria would likely increase the accuracy and precision of KG-BOOST with the trade-off of increasing the number of samples needed and, thus, the measurement time. The stopping criteria can be chosen based on the needed accuracy and precision, which may differ between applications of the method.

4.2 Gaussian processes in target optimization

This study also demonstrated that Gaussian process regression is suitable for modeling the MEP response function (*i.e.*, motor map). Gaussian processes allow taking into account the uncertainties of the problem, the biggest source of uncertainty being the MEP variability. Another advantage of Gaussian process regression is its suitability for modeling a response function of any smooth shape as opposed to, *e.g.*, parametric Gaussian curve fitting, which assumes that the underlying function is a symmetric Gaussian distribution as in Harquel *et al.* (2017). We consider nonparametric fitting of MEP curves advantageous, since the motor maps can be asymmetric (see example maps in Weiss *et al.* 2013, Julkunen 2014, and van de Ruit *et al.* 2015). Sampling in a dense grid, as we did, is beneficial, since it assists in revealing the shape and, thus, the location of the peak of the response curve better than sampling in a coarse grid. Furthermore, as the Gaussian process regression model links the data of neighboring points, the expected convergence speed of a sampling in a dense grid is similar to that of a sampling in a sparse grid.

When modeling with Gaussian processes, one needs to ensure that the responses are handled on a suitable scale. We chose a logarithmic scale to satisfy better the assumption of location/orientation-independent MEP variance included in the model. The BOOST algorithm seems to tolerate well the variability in the MEP variance that occurs in practice. Note that on a logarithmic scale, the peaks of the MEP response curves (Fig. 2) appear broader than on a linear scale. For ensuring a sufficient number of samples and for placing them optimally, we suggest adaptively guiding the sampling with, *e.g.*, knowledge gradient (as in KG-BOOST) or entropy-based methods (as implemented by Harquel *et al.* (2017)). There are also other methods that could suit for efficient guiding of the sampling, such as the expected-improvement method (Mockus *et al.* 1978) or the use of confidence bounds as sampling criteria (Cox and John 1992), but these were not tested in this study.

Even though Gaussian process regression is a nonparametric method, it includes several model parameters that need to be tuned case-specifically. Our suggestion for determining the parameters in the case of TMS-target optimization are presented in Section 2.1.2, but there are also other ways to determine these parameters. To adapt the posterior adequately to the data while avoiding overfitting, it was crucial to correctly tune the smoothness parameter $a_{1,d}$ (see Eq. (4)) to be of a reasonable magnitude. We chose to set this parameter with the simple zero-crossings formula and to keep it constant during the whole optimization procedure. Another option would be to tune $a_{1,d}$ among the other model parameters adaptively and to teach the model with the acquired data. For this purpose, we also tested maximum-likelihood estimation and cross-validation (see Chapter 5 of Rasmussen and Williams 2005). However, in our experience, these approaches tended to overestimate $a_{1,d}$, leading to overfitting the model to the data.

4.3 Future development and applications

Automated stimulation targeting could be further extended to adaptively adjust also the stimulation intensity (here, we used a predefined intensity, 110% rMT). In this case, the motor responses gathered during the target optimization might also be used in the rMT estimation, or the rMT could be determined separately after the target optimization with, *e.g.*, the adaptive algorithm presented by Awiszus (2003). In the future, the KG-BOOST algorithm can be applied for multi-dimensional problems, *e.g.*, to optimize simultaneously the location and orientation of the E-field, for example, with a 5-coil mTMS system similar to the one depicted by Koponen *et al*. (2018a). In principle, KG-BOOST can be implemented with any TMS system that allows automatic adjustment of the stimulation location and orientation, *e.g.*, with a robotically controlled stimulator. Furthermore, with a suitable software implementation, the algorithm could even guide the manual target search performed with a conventional navigated TMS system. With a multi-dimensional search space and a larger coverage of the cortex, one will be able to avoid the initial manual search, which we needed to place the transducer appropriately, and identify the optimal target even if it were situated abnormally.

The benefits of automated target optimization are applicable in several settings, from basic research to therapeutic uses of TMS. The automated target optimization could be used, *e.g.*, for studying the plasticity of motor cortex in a user-independent way. After modifying the function that guides the sampling, the BOOST algorithm may find applications in efficient and automated mapping of motor areas. If the actual cortical activation sites are of interest, one may combine the BOOST algorithm with individualized E-field modeling. In addition to finding the optimal stimulation parameters based on MEP amplitudes, the BOOST algorithm could be used to find an optimal stimulation target with respect to other available measures. For example, TMS targeting outside the motor cortex could be automatically optimized with respect to evoked cortical activity measured by electro-encephalography (Tremblay *et al*. 2019).

5 Conclusion

We demonstrated that electronically adjusted multi-locus TMS and Bayesian optimization provide a valid basis for automated search of TMS targets. The presented adaptively guided target search (KG-BOOST) gave results with good accuracy and precision, needing only a relatively small number of stimuli for convergence. We conclude that KG-BOOST enables fast, easy and user-independent target optimization, and that its benefits are applicable from basic research to therapeutic uses of TMS.

Acknowledgements

We thank Emma Skarstein for contributing to the literature search and investigating different aspects of modeling with Gaussian process regression. We thank Selja Vaalto and Ulf Ziemann for useful comments on the manuscript. This project has received funding from the Academy of Finland (Decisions No. 294625, 306845, and 327326), the Finnish Cultural Foundation, and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 810377).

Declaration of interests

RJI is an advisor and a minority shareholder of Nexstim Plc. The other authors declare no competing interests.

Appendix

Here, we present further details for computing the knowledge gradient. Assume that we have N noisy samples $\mathbf{y} = [y_1, \dots, y_N]^{\mathrm{T}}$ that correspond to the sampling parameters $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_N]$. Knowing them, we choose the next sampling parameters $\mathbf{x}_{N+1}$ with a knowledge-gradient sampling policy (Frazier *et al*. 2009, Frazier and Wang 2015). For getting the maximum of $\mathrm{KG}(\mathbf{x})$ (approximately), we choose an appropriately dense subset $\mathbf{X}^* = \{\mathbf{x}_m^*; m = 1, \dots, M\}$ in the search space. For a fixed $\mathbf{x}_m^*$, we form sequences $A(r)$ and $B(r)$, $r = 1, \dots, M$, by setting for each $\mathbf{x}_r^* \in \mathbf{X}^*$, $A(r) = \mu_N(\mathbf{x}_r^*)$ as in Eq. (6) and

$$B(r) = \frac{k(\mathbf{x}_r^*, \mathbf{x}_m^*) - \mathbf{k}(\mathbf{x}_r^*, \mathbf{X})(\mathbf{K} + \mathbf{\Lambda})^{-1}\mathbf{k}(\mathbf{x}_m^*, \mathbf{X})^{\mathrm{T}}}{[k(\mathbf{x}_m^*, \mathbf{x}_m^*) - \mathbf{k}(\mathbf{x}_m^*, \mathbf{X})(\mathbf{K} + \mathbf{\Lambda})^{-1}\mathbf{k}(\mathbf{x}_m^*, \mathbf{X})^{\mathrm{T}} + \lambda_*^2]^{1/2}},$$

where $\mathbf{k}(\mathbf{x}_r^*, \mathbf{X}) = [k(\mathbf{x}_r^*, \mathbf{x}_1), \dots, k(\mathbf{x}_r^*, \mathbf{x}_N)]$ and $\mathbf{k}(\mathbf{x}_m^*, \mathbf{X}) = [k(\mathbf{x}_m^*, \mathbf{x}_1), \dots, k(\mathbf{x}_m^*, \mathbf{x}_N)]$ contain the covariances between $\mathbf{x}_r^*$ or $\mathbf{x}_m^*$ and $\mathbf{x}_1, \dots, \mathbf{x}_N$ (formula for k in Eq. (4)), $\mathbf{\Lambda}$ is a diagonal matrix with $\lambda_1^2, \dots, \lambda_N^2$ on its diagonal, and λ_*^2 is the noise variance at location $\mathbf{x}_m^*$. As in Frazier *et al*. (2009), we choose subsequences $\alpha(s)$ and $\beta(s)$ of $A(r)$ and $B(r)$ and an additional sequence $\gamma(s)$, $s = 1, \dots, S$, by Matlab codes *AffinebreakpointsPrep.m* and *Affinebreakpoints.m* given in Frazier (2010). With these sequences, we get the knowledge-gradient function as

$$\mathrm{KG}(\mathbf{x}_m^*) = \sum_{s=1}^{S-1}\Big[\big(\beta(s+1) - \beta(s)\big)\Big(\varphi\big(\gamma(s)\big) - |\gamma(s)|\Phi(-|\gamma(s)|)\Big)\Big] + \max(\alpha(s)) - \mu_N^{\max},$$

where $\varphi(t) = (2\pi)^{-1/2}\exp(-t^2/2)$ and $\Phi(\mathrm{t}) = \int_{-\infty}^{t}\varphi(u)du$ are the probability density function and the cumulative density function of a normalized Gaussian random variable, respectively. Finally, we get the next sampling point as

$$\mathbf{x}_{N+1} = \arg\max_{\mathbf{x}_m^*} \mathrm{KG}(\mathbf{x}_m^*)\,.$$

References

Ahdab, R., Ayache, S.S., Brugières, P., Farhat, W.H., Lefaucheur, J.-P., 2016. The hand motor hotspot is not always located in the hand knob: A neuronavigated transcranial magnetic stimulation study. Brain Topogr. 29, 590–597. https://doi.org/10.1007/s10548-016-0486-2

Awiszus, F., 2003. TMS and threshold hunting. Suppl. Clin. Neurophysiol. 56, 13–23. https://doi.org/10.1016/S1567-424X(09)70205-3

Balslev, D., Braet, W., McAllister, C., Miall, R.C., 2007. Inter-individual variability in optimal current direction for transcranial magnetic stimulation of the motor cortex. J. Neurosci. Methods 162, 309–313. https://doi.org/10.1016/j.jneumeth.2007.01.021

Barker, A.T., Jalinous, R., Freeston, I.L., 1985. Non-invasive magnetic stimulation of human motor cortex. Lancet 325, 1106–1107. https://doi.org/10.1016/S0140-6736(85)92413-4

Bashir, S., Perez, J., Horvath, J.C., Pascual-Leone, A., 2013. Differentiation of motor cortical representation of hand muscles by navigated mapping of optimal TMS current directions in healthy subjects. J. Clin. Neurophysiol. 30, 390–395. https://doi.org/10.1097/WNP.0b013e31829dda6b

Bulubas, L., Sabih, J., Wohlschlaeger, A., Sollmann, N., Hauck, T., Ille, S., Ringel, F., Meyer, B., Krieg, S.M., 2016. Motor areas of the frontal cortex in patients with motor eloquent brain lesions. J. Neurosurg. 125, 1431–1442. https://doi.org/10.3171/2015.11.jns152103

Cox, D.D., John, S., 1992. A statistical method for global optimization, in: IEEE Conference on Systems, Man and Cybernetics. 1241–1246. https://doi.org/10.1109/ICSMC.1992.271617

Frazier, P., Powell, W., Dayanik, S., 2009. The knowledge-gradient policy for correlated normal beliefs. INFORMS J. Comput. 21, 599–613. https://doi.org/10.1287/ijoc.1080.0314

Frazier, P., 2010. MatlabKG library. Retrieved from http://people.orie.cornell.edu/pfrazier/src.html

Frazier, P.I., Wang, J., 2016. Bayesian optimization for materials design, in: Lookman, T., Alexander, F.J., Rajan, K. (Eds.), Information Science for Materials Discovery and Design. Springer, 45–75. https://doi.org/10.1007/978-3-319-23871-5_3

Harquel, S., Diard, J., Raffin, E., Passera, B., Dall'Igna, G., Marendaz, C., David, O., Chauvin, A., 2017. Automatized set-up procedure for transcranial magnetic stimulation protocols. NeuroImage 153, 307–318. https://doi.org/10.1016/j.neuroimage.2017.04.001

Julkunen, P., 2014. Methods for estimating cortical motor representation size and location in navigated transcranial magnetic stimulation. J. Neurosci. Methods 232, 125–133. https://doi.org/10.1016/j.jneumeth.2014.05.020

Kiers, L., Cros, D., Chiappa, K.H., Fang, J., 1993. Variability of motor potentials evoked by transcranial magnetic stimulation. Electroencephalogr. Clin. Neurophysiol. 89, 415–423. https://doi.org/10.1016/0168-5597(93)90115-6

Koponen, L.M., Nieminen, J.O., Ilmoniemi, R.J., 2018a. Multi-locus transcranial magnetic stimulation—theory and implementation. Brain Stimul. 11, 849–855. https://doi.org/10.1016/j.brs.2018.03.014

Koponen, L.M., Nieminen, J.O., Mutanen, T.P., Ilmoniemi, R.J., 2018b. Noninvasive extraction of microsecond-scale dynamics from human motor cortex. Hum. Brain Mapp. 39, 2405–2411. https://doi.org/10.1002/hbm.24010

Lefaucheur, J.-P., André-Obadia, N., Antal, A., Ayache, S.S., Baeken, C., Benninger, D.H., Cantello, R.M., Cincotta, M., de Carvalho, M., De Ridder, D., Devanne, H., Di Lazzaro, V., Filipović, S.R., Hummel, F.C., Jääskeläinen, S.K., Kimiskidis, V.K., Koch, G., Langguth, B., Nyffeler, T., Oliviero, A., Padberg, F., Poulet, E., Rossi, S., Rossini, P.M., Rothwell, J.C., Schönfeldt-Lecuona, C., Siebner, H.R., Slotema, C.W., Stagg, C.J., Valls-Sole, J., Ziemann, U., Paulus, W., Garcia-Larrea, L., 2014. Evidence-based guidelines on the therapeutic use of repetitive transcranial magnetic stimulation (rTMS). Clin. Neurophysiol. 125, 2150–2206. https://doi.org/10.1016/j.clinph.2014.05.021

Lefaucheur, J.-P., Picht, T., 2016. The value of preoperative functional cortical mapping using navigated TMS. Clin. Neurophysiol. 46, 125–133. https://doi.org/10.1016/j.neucli.2016.05.001

Lisanby, S.H., Luber, B., Perera, T., Sackeim, H.A., 2000. Transcranial magnetic stimulation: Applications in basic neuroscience and neuropsychopharmacology. Int. J. Neuropsychopharmacol. 3, 259–273. https://doi.org/10.1017/S1461145700002005

Meincke, J., Hewitt, M., Batsikadze, G., Liebetanz, D., 2016. Automated TMS hotspot-hunting using a closed loop threshold-based algorithm. NeuroImage 124, 509–517. https://doi.org/10.1016/j.neuroimage.2015.09.013

Mockus, J., Tiešis, V., Žilinskas, A., 1978. The application of Bayesian methods for seeking the extremum, in: Dixon, L.C.W., Szegö, G.P. (Eds.), Towards Global Optimization 2. North Holland Publ., 117–129.

Nieminen, J.O., Koponen, L.M., Mäkelä, N., Souza, V.H., Ilmoniemi, R.J., 2019. Short-interval intracortical inhibition in human primary motor cortex: A multi-locus transcranial magnetic stimulation study. NeuroImage 203, 116194. https://doi.org/10.1016/j.neuroimage.2019.116194

Oldfield, R.C., 1971. The assessment and analysis of handedness: The Edinburgh inventory. Neuropsychologia 9, 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

de Oliveira e Souza, V.H., 2018. Development of instrumentation for neuronavigation and transcranial magnetic stimulation. Doctoral dissertation, University of São Paulo.

Picheny, V., Wagner, T., Ginsbourger, D., 2013. A benchmark of kriging-based infill criteria for noisy optimization. Struct. Multidiscip. Optim. 48, 607–626. https://doi.org/10.1007/s00158-013-0919-4

Rasmussen, C.E., Williams, C.K.I., 2006. Gaussian processes for machine learning. The MIT Press.

Rossini, P.M., Barker, A.T., Berardelli, A., Caramia, M.D., Caruso, G., Cracco, R.Q., Dimitrijević, M.R., Hallett, M., Katayama, Y., Lücking, C.H., Maertens de Noordhout, A.L., Marsden, C.D., Murray, N.M.F., Rothwell, J.C., Swash, M., Tomberg, C., 1994. Non-invasive electrical and magnetic stimulation of the brain, spinal cord and roots: Basic principles and procedures for routine clinical application. Report from an IFCN committee. Electroencephalogr. Clin. Neurophysiol. 91, 79–92. https://doi.org/https://doi.org/10.1016/0013-4694(94)90029-9

Rossini, P.M., Burke, D., Chen, R., Cohen, L.G., Daskalakis, Z., Di Iorio, R., Di Lazzaro, V., Ferreri, F., Fitzgerald, P.B., George, M.S., Hallett, M., Lefaucheur, J.P., Langguth, B., Matsumoto, H., Miniussi, C., Nitsche, M.A., Pascual-Leone, A., Paulus, W., Rossi, S., Rothwell, J.C., Siebner, H.R., Ugawa, Y., Walsh, V., Ziemann, U., 2015. Non-invasive electrical and magnetic stimulation of the brain, spinal cord, roots and peripheral nerves: Basic principles and procedures for routine clinical and research application: An updated report from an I.F.C.N. Committee. Clin. Neurophysiol. 126, 1071–1107. https://doi.org/10.1016/j.clinph.2015.02.001

Scott, W., Frazier, P., Powell, W., 2011. The correlated knowledge gradient for simulation optimization of continuous parameters using Gaussian process regression. SIAM J. Optim. 21, 996–1026. https://doi.org/10.1137/100801275

Shahriari, B., Swersky, K., Wang, Z., Adams, R.P., de Freitas, N., 2016. Taking the human out of the loop: A review of Bayesian optimization. Proc. IEEE 104, 148–175. https://doi.org/10.1109/JPROC.2015.2494218

Tremblay, S., Rogasch, N. C., Premoli, I., Blumberger, D. M., Casarotto, S., Chen, R., Di Lazzarro, V., Farzan, F., Ferrarelli, F., Fitzgerald, P. B., Hui, J., Ilmoniemi, R. J., Kimiskidis, V. K., Kugiumtzis, D., Lioumis, P., Pascual-Leone, A., Pellicciari, M. C., Rajji, T., Thut, G., Zomorrodi, R., Ziemann, U., Daskalakis, Z. J., 2019. Clinical utility and prospective of TMS–EEG. Clinical Neurophysiology, 130, 802–844. https://doi.org/10.1016/j.clinph.2019.01.001

Valero-Cabré, A., Amengual, J.L., Stengel, C., Pascual-Leone, A., Coubard, O.A., 2017. Transcranial magnetic stimulation in basic and clinical neuroscience: A comprehensive review of fundamental principles and novel insights. Neurosci. Biobehav. Rev. 83, 381–404. https://doi.org/10.1016/j.neubiorev.2017.10.006

van de Ruit, M., Perenboom, M.J.L., Grey, M.J., 2015. TMS brain mapping in less than two minutes. Brain Stimul. 8, 231–239. https://doi.org/10.1016/j.brs.2014.10.020

Weiss, C., Nettekoven, C., Rehme, A.K., Neuschmelting, V., Eisenbeis, A., Goldbrunner, R., Grefkes, C., 2013. Mapping the hand, foot and face representations in the primary motor cortex — Retest reliability of neuronavigated TMS versus functional MRI. NeuroImage 66, 531–542. https://doi.org/10.1016/j.neuroimage.2012.10.046