Published in final edited form as:

IEEE Trans Neural Syst Rehabil Eng. 2019 June ; 27(6): 1320–1330. doi:10.1109/TNSRE.2019.2914475.

Optimal Estimation of Neural Recruitment Curves Using Fisher Information: Application to Transcranial Magnetic Stimulation

Seyed Mohammad Mahdi Alavi,

Faculty of Electrical Engineering, Shahid Beheshti University, Tehran, Iran. He was with the the Department of Psychiatry and Behavioral Sciences, Duke University, Durham, NC 27710, USA

Stefan M. Goetz [Member IEEE], and

Departments of Psychiatry and Behavioral Sciences, and Electrical and Computer Engineering, and Neurosurgery, Duke University, Durham, NC 27710, USA

Angel V. Peterchev [Senior Member IEEE]

Departments of Psychiatry and Behavioral Sciences, Biomedical Engineering, Electrical and Computer Engineering, and Neurosurgery, Duke University, Durham, NC 27710, USA

Abstract

This paper presents a novel method for fast and optimal determination of recruitment (input–output, IO) curve parameters in neural stimulation. A sequential parameter estimation (SPE) method was developed based on the Fisher information matrix (FIM), with a stopping rule based on successively satisfying a specified estimation tolerance. Simulated motor responses evoked by transcranial magnetic stimulation (TMS) were used as a test bed. Performance of FIM-SPE was characterized in 10 177 simulation runs for various IO parameter values corresponding to different virtual subjects, and compared with uniform sampling. Unlike uniform sampling, FIM-SPE identifies and samples the areas of the IO curve that contain maximum information about the curve parameters. For the most relaxed stopping rule, the median number of samples required for convergence was only 17 for FIM-SPE versus 294 for uniform sampling. For the highest reliability stopping rule, more than 92% of the FIM-SPE runs converged, with a median of 88 samples, whereas all uniform sampling runs reached 1000 samples without converging. Compared to uniform sampling, FIM-SPE reduced estimation errors up to 2-fold and required 10 times fewer stimuli. FIM-SPE could improve the speed and accuracy of determination of IO curves for neural stimulation. A software implementation of the algorithm is provided online.

Keywords

Neural stimulation; transcranial magnetic stimulation; input–output curve; recruitment curve; sequential parameter estimation; adaptive sample selection; Fisher information matrix; curve fitting; optimization

Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

m_alavi@sbu.ac.ir.

A software implementation in MATLAB of the algorithm and the examples presented in this paper are available online on Code Ocean: <https://doi.org/10.24433/CO.6ed10e0c-d749-4ba6-9b31-5f953d61acf8>

I. INTRODUCTION

In various neural stimulation applications, including in-vivo and ex-vivo neuronal characterization studies, implants for central or peripheral stimulation, functional muscular stimulation, as well as transcranial electric and magnetic stimulation, it is necessary to derive recruitment curves (or input–output (IO) curves) that link the strength of the stimulus to the resulting neural or neuromuscular activation. In this paper we focus on transcranial magnetic stimulation (TMS) as an example, but the IO curve characteristics and the proposed approach for their estimation are generalizable to other applications.

TMS IO curves are commonly determined for the primary motor cortex and comprise pairs of stimuli at various stimulation strengths and the amplitudes of corresponding motor responses, i.e., motor evoked potentials (MEPs) measured with electromyography (EMG) [1]. The IO curve enables a comprehensive noninvasive characterization of the neural stimulation target and is considered to capture the engagement of various neuronal populations [2]–[5]. The IO curve is used as a sensitive biomarker for the effect of medication or neuromodulatory interventions [6]–[8]. Whereas the IO curve by itself is a cloud of points, quantitative analysis requires the extraction of underlying system parameters. The average IO curve is generally sigmoidal in shape, and can therefore be characterized by parameters such as slope, midpoint, and low and high plateaus [9], [10]. However, there is a substantial stochastic component in the responses that can be modeled as variabilities at both the input and output of the sigmoid [3], [11]. As another feature of motor IO curves, the output magnitude and variability grow exponentially with stimulus intensity; therefore, MEP amplitude is preferably log-transformed before analysis [10], [12].

Conventional IO curve analysis involves the recording of a high number of stimulus–response pairs, which subsequently undergo regression with a parametrized sigmoidal function [2]. The regression is typically an offline process performed after the TMS data acquisition is completed without any feedback from the regression. Thus, conventional IO curve analysis cannot determine how much data should be acquired for reliable parameter estimation, and, consequently, the process is either unreliable or inefficient when too little or too much data are recorded, respectively. Conservative data sampling to ensure good parameter estimation can result in unnecessarily long data acquisition that increases the discomfort of the subject and introduces vulnerability to temporal drifts of the responses. In addition, not every sample increases the regression accuracy, as some stimulus amplitudes influence the regression outcome overproportionally [13].

This paper presents a novel method for fast and optimal estimation of neural stimulation IO curves using sequential parameter estimation (SPE) based on the Fisher information matrix (FIM). FIM-SPE adaptively determines the optimal intensity of each stimulus based on prior responses in order to maximize the extracted information about the IO parameters. This, in turn, minimizes the number of stimuli required to estimate the IO parameters, reducing the duration and discomfort of the procedure.

Sec. II describes the IO curve model. Sec. III provides a detailed description of the algorithm (the algorithm code is provided online to allow its adoption in practice). Sec. IV describes

the algorithm evaluation via simulations. The results are presented in Sec. V and discussed in Sections VI and VII.

II. NEURAL RESPONSE MODELS

Unlike prior work on parameter estimation of the probability of binary responses, for example for psychometric functions [14]–[16], the proposed method is for IO curves with continuous input and output variables. The example output signal considered here is the TMS-induced MER

A. Motor Evoked Potential Model for Testing

Since thorough characterization and comparison of the performance of parameter estimation algorithms requires a large amount of data both within and across subjects, in this paper we test the estimation methods with a large number of virtual subjects generated by an independent realistic model of MEP statistics [11]. The simulation test bed generates the peak-to-peak amplitudes of MEP responses as a function of TMS stimulus amplitude and includes individual features, such as activation threshold, maximum MEP size, sigmoidal MEP growth with increasing stimulation strength, and trial-to-trial variability (e.g. due to neural excitability fluctuations) according to a statistical representation of differences within and between healthy subjects [3], [12]. The model contains two Gaussian variabilities applied to the input and output of the sigmoid curve with standard deviations denoted by σ_x and σ_y , respectively [11]. Note that in these prior studies, the IO curve characteristics were extracted offline from a large dataset. In contrast, the objective of this paper is to develop efficient online estimation of individual IO curve parameters.

B. IO Curve Model for Parameter Estimation

A regression model used for online parameter estimation should fulfill several conditions. It should capture the parameters of most interest to scientists and clinicians, such as the IO curve midpoint, slope, and high saturation level (plateau). The model should allow a rapid regression within a reasonable number of samples. From a mathematical standpoint, the model needs to be globally identifiable, i.e., the model parameters can be estimated uniquely from the IO data [17], [18].

The sigmoid IO curve model [1], [2], [10] meets these requirements and is adopted in this paper. Our sigmoid IO model uses four parameters and is defined by

$$y = y^l + \frac{y^h - y^l}{1 + 10^{(m-x)s}}, \quad (1)$$

where x and y are the stimulus and response amplitudes, and y^l , y^h , m , and s denote the low plateau, high plateau, midpoint, and slope of the IO curve, respectively. Correspondingly, the parameter vector is defined as

$$\boldsymbol{\theta} = [\theta_1 \ \theta_2 \ \theta_3 \ \theta_4] = [y^l \ y^h \ m \ s]. \quad (2)$$

Throughout the paper, vectors and matrices are written in bold letters. For TMS, x and y are the amplitudes of the TMS pulse and MEP, respectively. The sigmoid model (1) is fitted to the log-transformed MEP data, since this approximately normalizes the MEP variability across the range of TMS stimulus amplitudes [3], [10], [12]. Finally, note that the proposed approach can also use sigmoid functions other than (1).

III. DESCRIPTION OF FIM-SPE

The objective is to estimate the IO curve model (1) and its parameter vector $\boldsymbol{\theta}$ by applying as few TMS stimuli as possible. This section details theoretical and practical aspects of the proposed method, which is based on sequential parameter estimation (SPE). SPE aims to identify parameters of a system by sequentially taking samples of the system and updating the estimation until a satisfactory level of accuracy is achieved [19]–[21]. Essentially, SPE is an advanced curve fitting approach, which determines the optimal location of samples in a sequential manner. Sample selection is also known as ‘optimal design’, ‘design of experiment’, or ‘sequential sampling’ [22]–[27]. With samples chosen optimally, the system parameters can be estimated with fewer experiments, which reduces the duration and cost of testing as well as the disruption of the system and associated risks it may cause. A typical SPE algorithm includes four steps, summarized in Fig. 1:

1. *Initialization*: A number of initial samples is taken from the system to start SPE.
2. *Estimation update*: A new estimate of the parameters is computed.
3. *Adaptive sample selection*: The optimal location of the next sample is computed, the sample is acquired, and the algorithm returns to *Estimation update*.
4. *Termination*: SPE terminates when a stopping rule is satisfied.

All samples are acquired within a feasible input range $x \in [x_{\min}, x_{\max}]$, where $x_{\min} = 0$ and $x_{\max} = 1$ correspond respectively to no pulse and maximum pulse amplitude, which is stimulator device dependent. If a subject cannot tolerate strong TMS pulses or for other reasons, $x_{\max}$ may be limited to less than 1. Set

$$\mathcal{D}_N = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}.$$

denotes the data after acquiring N samples, where (x_i, y_i) , $i = 1, 2, \dots, N$ is the i -th pair of stimulus amplitude and corresponding log-transformed MEP amplitude.

The estimation starts with an initial data set $\mathcal{D}_{N_0}$ where N_0 is the number of initial sample pairs. An initial estimation of $\boldsymbol{\theta}$, $\hat{\boldsymbol{\theta}}_{N_0}$, is obtained using curve fitting. The SPE method then computes the next stimulus amplitude, x_{N_0+1} , that should be applied to the subject. The

associated log-transformed MEP, y_{N_0+1} , is measured and the data set $\mathcal{D}_{N_0}$ is updated by adding the new pair (x_{N_0+1}, y_{N_0+1}) to it. By using the new data set $\mathcal{D}_{N_0+1}$, SPE obtains a new estimation of $\theta, \hat{\theta}_{N_0+1}$, and computes the next stimulus, x_{N_0+2} . This process is continued until the stopping rule is satisfied. The algorithm steps are elaborated below.

A. Initialization

The number of initial samples N_0 should be at least equal to the number of unknown parameters to achieve full rank of the Jacobian matrix in the curve fitting algorithm [28]. Selecting the location of the initial samples depends on the system nature, and is typically randomly or uniformly distributed [20], [21], [26]. We adopted an initialization strategy that increases efficiency of the IO curve estimation. Since the low plateau y^l is dominated by measurement noise, it can be estimated by samples from $x = 0$, i.e., baseline EMG data in the absence of a stimulus [10], [29]. Thus, the collection of the baseline data does not require simulating the subject or waiting between stimuli, and can therefore be acquired rapidly. The baseline data can be collected at any time during SPE, apart from the period of the stimulus evoked responses and their aftermath. In this paper, we collect n_{base} baseline data points once at the initialization stage. The baseline data are combined with n samples that are acquired for $x \neq 0$, e.g. MEPs acquired with TMS pulses, totaling

$$N = n_{\text{base}} + n$$

samples for curve fitting. Thus, n_{base} directly affects the speed of curve fitting.

With n_{base} baseline data, the rank of the Jacobian matrix will theoretically become more than four, and estimation of model (1) can be initialized. However, since parameters other than y^l are largely insensitive to the baseline data, it is suggested to add three other samples taken randomly from $x_{\min} < x < x_{\max}$, to provide initial information about y^h , m , and s . Hence, the characteristics of the initial data set $\mathcal{D}_{N_0}$ are

$$\begin{aligned} N_0 &= n_{\text{base}} + 3, \\ x_N &= 0, \text{ for } N = 1, \dots, n_{\text{base}}, \text{ and} \\ \text{random } x_N, & 0 < x_N \leq 1, N = N_0 - 2, N_0 - 1, N_0. \end{aligned}$$

B. Estimation Update

After every new sample, SPE updates the estimation of the parameter vector using a curve fitting algorithm, which minimizes a cost function. Thus, curve fitting is an iterative process that runs within each SPE step. A general form of the cost function used in curve fitting is

$$Q_N = \|y_N - \hat{y}_N\|_p^v + \lambda \|\theta - \hat{\theta}_N\|_p^v \quad (3)$$

where Q_N is the cost function given N samples, $\hat{y}_N$ is the estimate of the vector of observations y_N , $\hat{\theta}_N$ is the estimate of the parameter vector θ , and $\|\xi_N\|_p = \left(\sum_{i=1}^N |\xi_i|^p\right)^{1/p}$ denotes the p -norm of a vector $\xi_N = [\xi_1 \cdots \xi_N]$ [28], [30]. Parameter λ is used for a trade-off between the output and parameter estimation errors and depends on the curve fitting algorithm. There are numerous curve fitting methods which can be used for IO curve estimation, such as linear or nonlinear least squares, maximum likelihood, and Bayesian inference. The model characteristics, data, and desired specifications are important in the selection of the curve fitting algorithm [28], [30]. For instance, if a Bayesian inference method is used, prior statistical information on the parameter values is beneficial. In this paper, without loss of generality, we use a nonlinear least squares method with cost function (3), $p = v = 2$, and $\lambda = 0$, in combination with a trust-region optimization algorithm for the estimation update [31]–[33]. Using the trust-region algorithm, the curve fitting searches for the optimal estimate $\hat{\theta}_N$ between a lower bound $\hat{\theta}_{\min}$ and an upper bound $\hat{\theta}_{\max}$ [31]. We assume that $\hat{\theta}_{\min}$ and $\hat{\theta}_{\max}$ are fixed throughout the estimation process, although they can be adjusted after acquiring new samples. If no bound is defined on the parameters, Levenberg-Marquardt algorithm can also be used [34].

The low and high plateaus of the IO curve are on the order of microvolts and millivolts, respectively. As indicated above, the MEP amplitude is log-transformed before analysis. By adding some margins to these values, we set the lower limits of $\hat{y}^l$ and $\hat{y}^h$ to -7 and -4 and their upper limits to -3 and -1 in log space, respectively. We set the lower bound of $\hat{m}$ to 0.01 , since neural activation cannot be achieved with stimuli of zero intensity, and the upper limit of $\hat{m}$ to $x_{\max}$. If the slope parameter $s = 0$, then the IO curve becomes flat, which does not happen in reality; therefore $s > 0$. For large values of s (typically > 60 – 70), the range along the x -axis for the slope region would be small compared to the variability σ_x ; in this case the exact value of s does not affect much the IO curve. By considering some margins, we set the lower and upper limits of $\hat{s}$ to 0.01 and 100 , respectively. Thus, $\hat{\theta}_{\min}$ and $\hat{\theta}_{\max}$ for TMS IO curve estimation are set to

$$\begin{aligned}\hat{\theta}_{\min} &= [-7 \ -4 \ 0.01 \ 0.01]^T, \text{ and} \\ \hat{\theta}_{\max} &= [-3 \ -1 \ 1.00 \ 100]^T.\end{aligned}$$

The curve fitting requires an initialization, which is distinct from the SPE initialization. Every time that the curve fitting is run, the parameter vector is initialized, typically by a random-valued parameter vector, which converges to an estimation by minimizing the cost function (3) after some iterations. A prior knowledge of the parameter vector can be helpful in some cases for the curve fitting initialization. However, prior knowledge might require training data. Moreover, the accuracy of prior knowledge can be problematic and cause biased estimation. In this work, the trust-region curve fitting algorithm is assigned random initial parameter value guesses within $\hat{\theta}_{\min}$ and $\hat{\theta}_{\max}$.

There is always a possibility of a bad fit in all curve fitting methods. Bad fit refers to a divergence of the curve fitting algorithm, which is due to computational (or singularity) issues [35]. In many cases, a bad fit can be bypassed by repeating the curve fitting algorithm with a new random initialization. As a simple solution to detect and remove bad fits in each iteration, we compare the current estimate of the parameter with statistics of a number of previous estimates to identify unexpected large changes. If the difference is above a threshold value, the curve fitting command is repeated with a new random guess of the parameters. Thus, the set value of this threshold reflects a trade-off between computational effort and sensitivity to strongly fluctuating estimates which can be a sign of bad fits. As there is no natural rule for tuning this value, we explored it heuristically and identified a threshold value of 0.02 as an effective means to significantly reduce bad fits. Further, we apply the bad fit detection and removal after $n = 10$ samples (e.g., MEPs) have been acquired, to allow for initial settling of the estimates. See [35] for discussion of more advanced bad-fit detection and removal.

C. Sample Selection

After the initialization data have been acquired, SPE has to optimally choose the subsequent stimulus pulse strengths x to maximize the system information after each sample. The system information is typically quantified by using the gradient or likelihood functions. The Kullback-Leibler information measure [36] and stochastic approximation [37], [38] are examples of sequential sampling methods based on the gradient function maximization. This paper focuses on FIM, a quantitative metric of the system information which is based on the likelihood function [39]. It measures the sensitivity of likelihood functions to the system's parameters.

As mentioned earlier, the baseline data do not require stimulation and provide plenty of information for the estimation of y^l . Thus, the FIM is written in terms of

$$\tilde{\theta} = [y^h \ m \ s]$$

and the baseline data are excluded from the following computations. We define

$$\tilde{\mathbf{x}}_{N+1} = [x_{(n_{\text{base}}+1)} \ \cdots \ x_{N+1}]^T, \text{ and} \quad (4)$$

$$\tilde{\mathbf{y}}_{N+1} = [y_{(n_{\text{base}}+1)} \ \cdots \ y_{N+1}]^T \quad (5)$$

for $N \geq N_0$.

When x_{N+1} is applied, FIM with respect to $\tilde{\theta}$, $\tilde{\mathbf{x}}_{N+1}$, and $\tilde{\mathbf{y}}_{N+1}$ is given by

$$\mathbf{F}_{N+1} = \mathbb{E} \left[\left(\frac{\partial}{\partial \tilde{\boldsymbol{\theta}}} \ln \left(p(\tilde{\mathbf{y}}_{N+1} | \tilde{\mathbf{x}}_{N+1}, \tilde{\boldsymbol{\theta}}) \right) \right)^T \times \left(\frac{\partial}{\partial \tilde{\boldsymbol{\theta}}} \ln \left(p(\tilde{\mathbf{y}}_{N+1} | \tilde{\mathbf{x}}_{N+1}, \tilde{\boldsymbol{\theta}}) \right) \right) \right], \quad (6)$$

where p is the likelihood function and $N \geq N_0$.

With zero-mean normally distributed additive variability, FIM is given by [23]

$$\mathbf{F}_{N+1}(x_{N+1}, \tilde{\boldsymbol{\theta}}) = \frac{1}{\sigma_y^2} \tilde{\mathbf{F}}_{N+1},$$

where $\tilde{\mathbf{F}}_{N+1} = [\tilde{f}_{N+1}(q, r)]$, $q, r = 1, \dots, 3$, and

$$\tilde{f}_{N+1}(q, r) = \sum_{i=n_{\text{base}}+1}^{N+1} \frac{\partial y(x_i | \tilde{\boldsymbol{\theta}})}{\partial \tilde{\theta}_q} \frac{\partial y(x_i | \tilde{\boldsymbol{\theta}})}{\partial \tilde{\theta}_r}. \quad (7)$$

$\tilde{\mathbf{F}}_{N+1}$ is recursively computed per [40]

$$\tilde{\mathbf{F}}_{N+1}(x_{N+1}, \tilde{\boldsymbol{\theta}}) = \tilde{\mathbf{F}}_N(\tilde{\mathbf{x}}_N, \tilde{\boldsymbol{\theta}}) + \delta \tilde{\mathbf{F}}_{N+1}(x_{N+1}, \tilde{\boldsymbol{\theta}}) \quad (8)$$

where $\delta \tilde{\mathbf{F}}_{N+1}(x_{N+1}, \tilde{\boldsymbol{\theta}}) = [\delta \tilde{f}_{N+1}(q, r)]$ and

$$\delta \tilde{f}_{N+1}(q, r) = \frac{\partial y(x_{N+1}, \tilde{\boldsymbol{\theta}})}{\partial \tilde{\theta}_q} \frac{\partial y(x_{N+1}, \tilde{\boldsymbol{\theta}})}{\partial \tilde{\theta}_r}. \quad (9)$$

A property of FIM is the Cramér-Rao bound,

$$\text{cov}(\hat{\boldsymbol{\theta}}_N) \geq \tilde{\mathbf{F}}_N^{-1}(\tilde{\boldsymbol{\theta}}), \quad N \geq N_0 \quad (10)$$

where cov denotes covariance, $\tilde{\mathbf{F}}^{-1}$ is the FIM inverse, and $\hat{\boldsymbol{\theta}}$ is the estimate of the parameter vector $\tilde{\boldsymbol{\theta}}$ [41], [42]. Thus, the larger the FIM, the smaller the lower bound of the covariance matrix, which is equivalent to a better estimation. In the case of a single-parameter estimation problem, FIM is a real-valued function that can be maximized through numerical optimization, resulting in minimum variance of estimates, which guarantees convergence for unbiased estimators (see [43] for an example). However, maximization of a matrix, such as in the multi-parameter case of neural recruitment curves, is not defined in general. Instead, some real-valued FIM summary metric (optimality criterion) is maximized. A widely used criterion is D-optimality, which maximizes the determinant of FIM [44], [45]. Other criteria include E- and ED-optimality which maximize the FIM minimum eigenvalue and

determinant expectation, respectively [46]. In this work, we use D-optimality since, in addition to the information on the parameters separately, it considers the joint information (covariation) among the parameters, which supports estimation quality with respect to all parameters [47]. Extension to other optimality criteria is straightforward.

In the D-optimal framework, each sample location x_{N+1} is obtained by solving the optimization problem

$$\begin{aligned} \underset{x_{N+1}}{\text{minimize}} \quad & -\det(\tilde{\mathbf{F}}_{N+1}(x_{N+1}(x_{N+1}, \hat{\boldsymbol{\theta}}_N)) \\ \text{subject to} \quad & x_{\min} \leq x_{N+1} \leq x_{\max} \\ & N \geq N_0. \end{aligned} \quad (11)$$

Further information about the computation of FIM and Cramér-Rao bound can be found in [48]–[51].

Algorithm 1 summarizes the computational steps in finding the next optimal sample in IO curve estimation. Note that the D-optimal problem (11) is a non-convex optimization problem in general [52], which demands nonlinear programming tools for finding the global optimum. Convexity of the D-optimal problem for a specific linear system/model is studied in §7.5 of [53]. MATLAB's 'ga', 'GlobalSearch', or 'MultiStart' from the genetic algorithm and global optimization toolboxes are nonlinear programming algorithms that can be used to solve the D-optimal problem (11). GlobalSearch is used in this work.

D. Termination

We define a convergence criterion as

$$\left| \frac{\hat{\theta}_{l,N} - \hat{\theta}_{l,N-1}}{\hat{\theta}_{l,N-1}} \right| < \epsilon_l \text{ for } l = 1, \dots, 4 \text{ and } N \geq N_0 \quad (12)$$

where ϵ_l denotes the estimation tolerance of θ_l . Condition (12) does not require any knowledge about the parameters' range or order of magnitude. There is no specific rule for selecting ϵ_l . A large ϵ_l can lead to termination with a high estimation error, whereas a very small ϵ_l will not necessarily improve the estimation performance. An estimation tolerance of $\epsilon_l = 0.001$ for $l = 1, \dots, 4$ is used in this work.

The SPE can be terminated if the convergence criterion (12) is satisfied T successive times for all parameters. This is referred to as the *stopping rule*. Although larger numbers of samples may be required, simulations show that multiple successive satisfaction of (12) (i.e., $T > 1$) improves the estimation performance. Similar to ϵ_l , it is unclear how to choose the value of T in the stopping rule [54]. The size of the parameter vector, model complexity, and/or noise characteristics may affect the best choice for T . We therefore study the effect of

$T = 1, \dots, 5$ on the simulation results. If the estimation has fulfilled the stopping rule at $N = N_f$, the number of stimulus pulses at which termination occurs is $n_f = N_f - n_{\text{base}}$.

SPE is also terminated if $n = n_{\text{max}}$, where n_{max} is the maximum allowable number of stimuli. In this work, we set $n_{\text{max}} = 1000$, which is representative of a practical limit of stimulation.

IV. METHOD OF FIM-SPE TESTING

The proposed FIM-SPE method for IO curve estimation was simulated 10 177 times for various IO parameter values corresponding to different, randomly selected virtual subjects. MEP responses were generated with the realistic model described in Sec. II. Fig. 2(a) displays 10000 uniformly-distributed stimulus–MEP pairs from a representative virtual subject. The IO curve model (1) was fitted to these data to provide ‘ground truth’ reference values of parameters θ . Parameters $\hat{\theta}$ were then estimated using the FIM-SPE algorithm and compared against the reference values. The baseline data comprised $n_{\text{base}} = 100$ samples, and the input range was set to $x_{\text{min}} = 0$ and $x_{\text{max}} = 1$.

Every time a sample was taken and the estimation updated, the result was compared with *uniform sampling*, which uses the same number of samples but uniformly distributed along x . Uniform sampling exemplifies an algorithm that does not strategically choose the placement of the samples, and is therefore theoretically suboptimal. Uniform sampling or variations thereof are commonly used in neural IO curve estimation. In TMS, the range of sampling is often constrained by first determining a motor threshold that estimates the lower reach of the IO slope; uniform or quasi-uniform sampling is then applied to the guessed extent of the slope and sometimes the high plateau region [1], [2], [5], [10]. We did not model such elaborate schemes since they are generally ad hoc, vary across studies, and involve two stages of estimation—the motor threshold is a combination of the IO midpoint and slope parameters and can be estimated in various ways, often suboptimal in practice [7], [55]–[57]. Moreover, if motor threshold data are available, they can be used to initialize FIM-SPE as well. Therefore, we focused on uniform sampling as a simple, well defined benchmark. To ensure a fair comparison, the same initialization data were used for uniform sampling as in FIM-SPE.

To quantify the algorithm performance, we calculated the absolute relative error and relative bias of the estimation. After acquiring N samples and updating the estimation with $\mathcal{D}_N$, the relative error associated with parameter θ_l was computed as

$$e\theta_{l,N} = \frac{\hat{\theta}_{l,N} - \theta_l}{\theta_l}, \quad l = 1, \dots, 4. \quad (13)$$

The absolute relative error of estimation was computed by taking the absolute value of (13). After normalization of the distribution [58], [59], these errors were compared statistically using a mixed model (JMP Pro 14, SAS), with sampling method, stopping rule parameter T , and their interaction as fixed effects, followed by post-hoc Tukey difference tests. Finally,

the relative estimation bias for each parameter was computed by taking the expected value of the relative error (13) over all simulation runs.

V. RESULTS

A. Representative Parameter Estimation Run

Figures 2 and 3 show the estimation results for a representative simulation run. FIM-SPE converged and terminated faster than uniform sampling for all values of stopping rule parameter T . For $T = \{1, 2, 3, 4, 5\}$, FIM-SPE terminated at $n_f = \{13, 14, 15, 103, 104\}$, respectively. In contrast, the uniform sampling estimation method only satisfied the stopping rule for $T = 1$ with $n_f = 229$, and did not terminate for larger T values up to $n_{\max} = 1000$ samples.

Fig. 2(b)–(c) shows that FIM-SPE focuses on three clusters of x values, which provide maximum information about the midpoint, slope, and high plateau of the IO curve. There is little variability in the x -axis locations where FIM-SPE samples the curve, which indicates robustness of the method with respect to x and y variability. In this FIM-SPE implementation, MEPs in the low plateau are not sampled, since this information is obtained from the baseline EMG data and parameter y^l is excluded from the FIM.

Fig. 3 shows the estimation steps of the IO curve parameters versus the number of samples up to $n = 104$, when FIM-SPE satisfied the stopping rule with $T = 5$. The estimation of the low plateau was instantaneous due to the availability of baseline EMG data. The high plateau was estimated relatively fast by sampling at the maximum device output (if needed for tolerability, the maximum TMS strength can be limited to lower levels). The midpoint took longer since it requires sampling in the IO slope region where the variability is higher. Finally, the slope took the longest to estimate. The likely reason is that, in addition to the high variability in the slope region, it is technically a derivative estimate and thus needs information from at least two x points, unlike the other parameters. This is also demonstrated in Fig. 4, which shows the evaluation of the convergence criterion (12) for termination.

B. Number of Samples for Estimation

Fig. 5 shows the percentage of runs that satisfied the stopping rule for FIM-SPE and uniform sampling. The larger the value of T , the more samples n_f it took for the algorithms to terminate. FIM-SPE significantly outperformed uniform sampling in terms of the probability of satisfying the stopping rule for $n_f \leq 1000$. In the worst case, with $T = 5$, 92% of runs using FIM-SPE successfully terminated. In contrast, none of the runs could satisfy the stopping rule with uniform sampling for $T = 4$ or $T = 5$, and few did for $T = 2$ or $T = 3$.

Fig. 6 shows the statistics of samples-to-termination n_f for FIM-SPE and uniform sampling versus parameter T in the stopping rule. FIM-SPE achieved markedly smaller n_f compared to uniform sampling. The median of required samples in FIM-SPE method was 17 for $T = 1$ and increased to 88 for $T = 5$. The required number of samples was dramatically larger for uniform sampling, with median of 294 for $T = 1$ and many runs for larger values of T reaching $n_{\max} = 1000$ without terminating.

C. Estimation Error

For runs where FIM-SPE converged, Fig. 7 summarizes the absolute relative errors of estimations for FIM-SPE and uniform sampling. Compared to the other three parameters, the low plateau was estimated most accurately, with median error of $< 0.18\%$ for both FIM-SPE and uniform sampling. The slope estimation had the largest error by far, reaching a median value of nearly 50% and 95th percentile value of approximately 800% for uniform sampling.

FIM-SPE improved the estimation error of the slope, midpoint, and high plateau ($F(1,95066) > 190$, $p < 0.0001$) with absolute reductions of 8.73%, 1.56%, and 0.7% in the error median values, respectively. Moreover, the 95th percentile values for FIM-SPE were lower than for uniform sampling up to 2-fold. There was no significant main effect of sampling method on the estimation error of the low plateau, but there was a significant interaction ($F(4,95066) = 17.3$, $p < 0.0001$) corresponding to slightly lower error for uniform sampling than for FIM-SPE for $T = \{3, 4, 5\}$. This is expected since uniform sampling collects more samples from the low plateau region; however, it is trivial to add more baseline EMG data to reduce the error of FIM-SPE further, if needed.

By increasing T , the errors in all parameters decreased ($F(4,95066) > 23$, $p < 0.0001$), with most pronounced effects for the midpoint and slope. For FIM-SPE and uniform sampling, the median estimation error of the slope was 40.2% and 48.9% for $T = 1$ and decreased to 15.5% and 18.7% for $T = 5$, respectively.

The estimation bias of the slope s also decreased monotonically with increasing T from 92.7% to 17.6% for FIM-SPE and from 156% to 34.9% for uniform sampling. The estimation bias of the slope parameter s was positive, i.e., on average a steeper IO curve was estimated. For the other parameters, the bias was small, not exceeding 2% in absolute value for all values of T for both FIM-SPE and uniform sampling.

D. Effect of IO Parameters on Number of Samples

Fig. 8 illustrates the effect of IO parameters y^l , y^h , m , s , σ_x , and σ_y on the number of samples required to reach convergence n_f . The data for only $T = 5$ are shown, since the results for $T = 1, \dots, 4$ were similar.

Parameters s and σ_x had the strongest (positive) correlations with n_f . This is expected since larger s means sharper slope, which makes sample selection in the slope area challenging. The larger σ_x affects the output variability nonlinearly with strongest effect in the slope area, making estimation more challenging. The specific values of y^l and y^h do not qualitatively change the estimation difficulty, hence the lack of correlation with n_f . Variability σ_y had only a minor effect, likely because its contribution to the MEP amplitude averages out to zero over multiple samples.

VI. DISCUSSION

FIM-SPE identifies and samples the regions of the IO curve with highest information for parameter estimation. Our simulations demonstrate that this approach has substantially

better estimation speed and error compared to uniform sampling. The results further show that the estimation error appears more stable for FIM-SPE as it generated fewer outlier error values compared to uniform sampling.

Similar to reports in the literature [9], the error and bias in estimating the IO curve slope were relatively large. However, in contrast to the observation that adaptive sampling tends to increase the estimation bias of the slope compared to uniform sampling [9], the error and bias were both lower for FIM-SPE than uniform sampling. Furthermore, the error and bias decreased monotonically with increasing value of the termination rule parameter T . The relatively large estimation error and bias of the slope parameter are likely due to the fact that information from at least two distinct x-axis points is necessary to compute the slope and that the slope area exhibits the largest variability and heteroscedasticity due to the nonlinear contribution of input-side (x-axis) variability [3]. This is also probably why the slope parameter critically influences the number of samples to termination, as seen in Fig. 8. Approaches to reduce this error can be explored in the future, potentially including more emphasis on information extraction from the slope area, a different IO curve function, and explicit modeling of input-side variability. Further, the parameter error quantification in our study was confounded since ground-truth parameter values were not available, because the MEP data were generated by a model more complex than the fitted IO curve, and the input-side variability has nonlinear impact on the output precluding easy computation of parameter means. Our approach of approximating the true parameter values with reference fits of a very large number of uniform samples can contribute to the computed errors.

In the implemented FIM-SPE, the low plateau parameter is fitted after each new FIM-SPE sample, even though the algorithm does not sample this regions due to the abundance of baseline EMG data. As an alternative approach, the low plateau could be estimated once in the beginning of the procedure using the baseline data, and fixed until termination. Thus, only a three-parameter curve fitting is run for each estimation update, speeding up the computations. In a smaller side study (data not shown), we tested this alternative approach and observed the same performance trends, with slightly reduced number of samples to termination but somewhat higher parameter estimation errors compared to the reported implementation.

Other aspects of the algorithm can be developed further as well. The computation of the FIM in the presence of input-side variability requires more research [3]. Extending the algorithm to estimate the input- and output-side IO variability can provide more quantification and insight of the underlying neurophysiology [3]. These can also be enhanced by providing the option of multiple input parameters, e.g. pulse width in addition to amplitude [10]. FIM-SPE can be adapted to stimulus-response curves that are non-sigmoidal. Of interest are also the study of other optimality criteria and curve fitting algorithms, such as maximum likelihood and Bayesian. For example, the Bayesian maximum a posteriori probability estimator has been employed for curve fitting [40]. This estimator was reported to improve the estimation error up to 3%. However, the sample selection was ‘critically’ dependent on prior information about the parameters, which may require a large database of individual IO data.

VII. CONCLUSION

FIM-SPE is efficient and practical for IO curve parameter estimation, and outperforms uniform sampling by more than 10-fold in speed and up to 2-fold in accuracy. FIM-SPE can improve the speed and accuracy of IO curve determination in TMS and other in-vivo and ex-vivo neural stimulation applications. A software implementation of the algorithm is provided online to facilitate its adoption and further development.

Acknowledgements

This work was supported by the National Institutes of Health under grant nos. R01MH091083, R01NS088674, and RF1MH114268; the Brain & Behavior Research Foundation under NARSAD award no. 22796; and by Shahid Beheshti university. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.

REFERENCES

- [1]. Devanne H et al., "Input-output properties and gain changes in the human corticospinal pathway," *Exp. Brain Res*, vol. 114, no. 2, 329–338, 1997. [PubMed: 9166922]
- [2]. Pitcher JB et al., "Age and sex differences in human motor cortex input-output characteristics," *J. Physiol*, vol. 546, no. 2, pp. 605–613, 2003. [PubMed: 12527746]
- [3]. Goetz SM et al., "A novel model incorporating two variability sources for describing motor evoked potentials," *Brain Stimul*, vol. 7, no. 4, pp. 541–552, 2014. [PubMed: 24794287]
- [4]. Di Lazzaro V et al., "I-wave origin and modulation," *Brain Stimul*, vol. 5, no. 4, 512–525, 2012. [PubMed: 21962980]
- [5]. D'Ostilio K et al., "Effect of coil orientation on strength-duration time constant and I-wave activation with controllable pulse parameter transcranial magnetic stimulation," *Clin. Neurophysiol*, 127(1), pp. 675–683, 2016. [PubMed: 26077634]
- [6]. Gangitano M et al., "Modulation of input-output curves by low and high frequency repetitive transcranial magnetic stimulation of the motor cortex," *Clin. Neurophysiol*, 113(8), pp. 1249–1257, 2002. [PubMed: 12140004]
- [7]. Rossini PM et al., "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(6), pp. 1071–1107, 2015. [PubMed: 25797650]
- [8]. Ziemann U et al., "TMS and drugs revisited," *Clin. Neurophysiol*, 126(10), pp. 1847–68, 2015 [PubMed: 25534482]
- [9]. Kaernbach C, "Slope bias of psychometric functions derived from adaptive data," *Perception & Psychophysics*, 63(8), pp. 1389–1398, 2001. [PubMed: 11800464]
- [10]. Peterchev AV et al., "Pulse width dependence of motor threshold and input-output curve characterized with controllable pulse parameter transcranial magnetic stimulation," *Clin. Neurophysiol*, vol. 124, no. 7, pp. 1364–1372, 2013. [PubMed: 23434439]
- [11]. Goetz SM et al., "Statistical model of motor evoked potentials for simulation of transcranial magnetic and electric stimulation," *bioRxiv*, 406777, 2018. doi:10.1101/406777
- [12]. Nielsen JF, "Logarithmic distribution of amplitudes of compound muscle action potentials evoked by transcranial magnetic stimulation," *J. Clin. Neurophysiol*, vol. 13, no. 5, 423–434, 1996. [PubMed: 8897207]
- [13]. Treutwein B, "Adaptive psychophysical procedures," *Vision Res*, vol. 35, no. 17, pp. 2503–2522, 1995. [PubMed: 8594817]
- [14]. Green DM, "Stimulus selection in adaptive psychophysical procedures," *J. Acoust. Soc. Am*, vol. 87, pp. 2662–2674, 1990. [PubMed: 2373801]
- [15]. Kontsevich LL and Tyler CW, "Bayesian adaptive estimation of psychometric slope and threshold," *Vision Res*, vol. 39, pp. 2729–2737, 1999. [PubMed: 10492833]

- [16]. Shen Y and Richards VM, "A maximum-likelihood procedure for estimating psychometric functions: Thresholds, slopes, and lapses of attention," *J. Acoust. Soc. Am*, vol. 132, pp. 957–967, 2012. [PubMed: 22894217]
- [17]. Raue A et al., "Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood," *Bioinformatics*, vol. 25, pp. 1923–1929, 2009. [PubMed: 19505944]
- [18]. Alavi SMM et al., "Identifiability of Generalized Randles Circuit Models," *IEEE Trans. Control Syst. Technol*, Vol. 25, No. 6, 2112–2120, 2017.
- [19]. Ford I and Silvey SD, "A sequentially constructed design for estimating a nonlinear parametric function," *Biometrika*, vol. 67, no. 2, pp. 381–388, 1980.
- [20]. Chang YC and Martinsek AT, "Fixed size confidence regions for parameters of a logistic regression model," *Ann. Statist*, vol. 20, pp. 1953–1969, 1992.
- [21]. Chang YC and Lu HY, "Online calibration via variable length computerized adaptive testing," *Psychometrika*, vol. 75, pp. 140–157, 2010.
- [22]. Fedorov VV, *Theory of optimal experiments*, Elsevier, 1972.
- [23]. Goodwin GC and Payne RL, *Dynamic system identification: experiment design and data analysis*, Academic press, 1977.
- [24]. Atkinson AC, "Developments in the design of experiments," *International Statistical Review*, vol. 50, pp. 161–177, 1982.
- [25]. Pukelsheim F, *Optimal design of experiments*, ser. Probability and mathematical statistic, Wiley, 1993.
- [26]. Berger MPF, "Sequential sampling designs for the two parameter item response theory model," *Psychometrika*, vol. 57, no. 4, pp. 521–538, 1992.
- [27]. Pronzato L, "Optimal experimental design and some related control problems," *Automatica*, vol. 44, no. 2, pp. 303–325, 2008.
- [28]. Bertsekas DP, *Nonlinear programming*, Athena scientific, 1999.
- [29]. Moeller C et al., "Hysteresis effects on the input–output curve of motor evoked potentials," *Clin. Neurophysiol*, vol. 120, no. 5, 1003–1008, 2009. [PubMed: 19329358]
- [30]. Millar RB, *Maximum likelihood estimation and inference: with examples in R, SAS and ADMB*, John Wiley & Sons, 2011.
- [31]. *Curve Fitting Toolbox User's Guide*. The MathWorks Inc., 2016.
- [32]. Byrd RH et al., "Approximate Solution of the Trust Region Problem by Minimization over Two-Dimensional Subspaces," *Mathematical Programming*, vol. 40, no. 1–3, pp. 247–263, 1988.
- [33]. Sorensen DC, "Newton's method with a model trust region modification," *SIAM J. Numer. Anal*, 19(2), pp. 409–426, 1982.
- [34]. Moré JJ, "The Levenberg-Marquardt algorithm: implementation and theory," in *Numerical analysis*, Springer, 1978.
- [35]. Nguyen T et al., "Robust dose-response curve estimation applied to high content screening data analysis," *Source Code Biol. Med*, vol. 9, pp. 1–9, 2014. [PubMed: 24401704]
- [36]. Weinstein E et al., "Sequential algorithms for parameter estimation based on the kullback-leibler information measure," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 9, pp. 1652–1654, 1990.
- [37]. Robbins H and Monro S, "A stochastic approximation method," *Ann. Math. Statist*, vol. 22, no. 3, pp. 400–407, 1951.
- [38]. Kiefer J and Wolfowitz J, "Stochastic estimation of the maximum of a regression function," *Ann. Math. Statist*, vol. 23, no. 3, pp. 462–466, 1952.
- [39]. Fisher RA, *The Design of Experiments* (9th ed.), Macmillan, 1971 (1935 1st ed.).
- [40]. Santos N et al., "Optimal sampling and estimation in PASL perfusion imaging," *IEEE Trans. Biomed. Eng*, vol. 58, no. 11, pp. 3165–3174, 2011. [PubMed: 21846602]
- [41]. Cramér H, *Mathematical Methods of Statistics*, Princeton Univ. Press Princeton, NJ, 1946.
- [42]. Rao CR, "Information and the accuracy attainable in the estimation of statistical parameters," *Bulletin of the Calcutta Mathematical Society*, vol. 37, pp. 81–89, 1945.

- [43]. Yilmaz Y and Wang X, "Sequential decentralized parameter estimation under randomly observed Fisher information," *IEEE Trans. Inf. Theory*, vol. 6, no. 2, pp. 1281–1300, 2014.
- [44]. De Cock A et al., "D-optimal input design for nonlinear FIR-type systems: A dispersion based approach," *Automatica*, vol. 73, pp. 88–100, 2016.
- [45]. Rothenberger M. Maximizing Parameter Identifiability of an Equivalent Circuit Battery Model Using Optimal Periodic Input Shaping; ASME 2014 Dynamic Systems and Control Conference; 2014. 1–10.
- [46]. Atkinson A et al., *Optimum Experimental Designs, With SAS*, Oxford University Press, 2007.
- [47]. St.John RC and Draper NR, "D-optimality for regression designs: a review," *Technometrics*, vol. 17, no. 1, pp. 15–23, 1975.
- [48]. Pakrooh P et al., "Analysis of Fisher information and the Cramér–Rao bound for nonlinear parameter estimation after random compression," *IEEE Trans. Signal Process*, vol. 63, no. 23, pp. 6423–6428, 2015.
- [49]. Gini F et al., "The modified cramer-rao bound in vector parameter estimation," *IEEE Trans. Commun*, vol. 46, no. 1, pp. 52–60, 1998.
- [50]. Stoica P et al., "Cisoid parameter estimation in the colored noise case: asymptotic cramer-rao bound, maximum likelihood, and nonlinear least-squares," *IEEE Trans. Signal Process*, vol. 45, no. 8, pp. 2048–2059, 1997.
- [51]. Ober RJ et al., "Calculation of the Fisher information matrix for multidimensional data sets," *IEEE Trans. Signal Process*, vol. 51, no. 10, pp. 2679–2691, 2003.
- [52]. Liu Y et al., "D-optimal design for parameter estimation in discrete-time nonlinear dynamic systems," *Mathematical Problems in Engineering*, vol. 2012, Article ID 296701, pp. 1–13, 2012.
- [53]. Boyd S and Vandenberghe L, *Convex Optimization*, Cambridge University Press, 2004.
- [54]. Spall JC, *Introduction to Stochastic Search and Optimization: Estimation, Simulation, and Control*, Wiley, 2003.
- [55]. Awiszus F, "TMS and threshold hunting," In: Paulus W, Tergau F, Nitsche MA, Rothwell JC, Ziemann U, Hallett M, eds *Transcranial Magnetic Stimulation and Transcranial Direct Current Stimulation (Supplements to Clin. Neurophysiol.)*, vol. 56, pp. 13–23, Elsevier Science, 2003.
- [56]. Götz S et al., "Threshold estimation with transcranial magnetic stimulation: algorithm comparison," *Clin. Neurophysiol*, vol. 122, pp. S197, 2011.
- [57]. Qi F et al., "Fast estimation of transcranial magnetic stimulation motor threshold," *Brain. Stimul.* vol. 4, pp. 50–7, 2011. [PubMed: 21255756]
- [58]. Seiler FA and Alvarez JL, "On the selection of distributions for stochastic variables," *Risk Analysis*, vol. 16, no. 1, 5–18, 1996.
- [59]. Gorla MN, "A normal approximation for the chi-square distribution," *Australian and New Zealand Journal of Statistics*, vol. 34, no. 1, 55–64, 1992.

Fig. 1:
Flowchart of sequential parameter estimation (SPE).

Fig. 2:

Example simulated TMS MEP data and corresponding IO curve fits for (a) reference data of $n = 10,000$ samples and FIM-SPE and uniform sampling for (b) $n = 104$ and (c) $n = 1000$. x-axis is normalized TMS pulse amplitude and y-axis is MEP peak-to-peak amplitude.

Fig. 3:

Estimated IO curve parameter values during FIM-SPE and uniform sampling in Fig. 2(b): (a) low plateau y^l , (b) high plateau y^h , (c) midpoint m , and (d) slope s . x-axis shows number of stimuli n . Estimation sequence comprises $n = 104$ samples for which FIM-SPE satisfies stopping rule with $T = 5$.

Fig. 4:
Fulfillment of convergence criterion for termination of parameter estimation with FIM-SPE for the example run in Fig. 2(b). The y-axis was scaled for a better visualization.

Fig. 5:
Percentage of runs that satisfy the stopping rule as a function of parameter T .

Fig. 6:

Samples-to-termination, n_f as a function of stopping rule parameter T . Markers and whiskers show median and 5th–95th percentile range, respectively.

Fig. 7:

Absolute relative errors for FIM-SPE and uniform sampling at the sample when FIM-SPE converged, n_f for various values of T in the stopping rule and parameters (a) low plateau y^l , (b) high plateau y^h , (c) midpoint m , and (d) slope s . Markers and whiskers show median and 5th–95th percentile values, respectively. The results for the two estimation algorithms are offset on the x-axis for better visualization. Asterisk (*) indicates significantly lower error than alternative sampling method ($p < 0.0001$).

Fig. 8:

Correlation of number of samples before termination, n_f of FIM-SPE with $T=5$ and MEP model parameters (a) low plateau y^l , (b) high plateau y^h , (c) midpoint m , (d) slope s , (e) y-axis variability σ_y , and (f) x-axis variability σ_x for each simulation run. Included in each plot are linear regression trend lines, correlation coefficient r , and significance value p .

Algorithm 1

Selection of x_{N+1} from most recent data set $\mathcal{D}_N$ and parameter estimate $\hat{\theta}_N$

1: Compute recursively $\tilde{\mathbf{F}}_{N+1}$ for $N = N_0$ by using (8). The sensitivity functions $\partial y(x_i, \hat{\theta}_N) / \partial \hat{\theta}_r$, $r = 1, 2, 3$ are obtained as follows:

$$g_{i,N}(x_i, \hat{\theta}_N) = \frac{1}{(\hat{\theta}_{3,N} - x_i) \hat{\theta}_{4,N} + 10}$$

$$\frac{\partial y(x_i, \hat{\theta}_N)}{\partial \hat{\theta}_1} = g_{i,N}(x_i, \hat{\theta}_N)$$

$$\frac{\partial y(x_i, \hat{\theta}_N)}{\partial \hat{\theta}_2} = - \frac{\ln(10) \hat{\theta}_{4,N} (\hat{\theta}_{2,N} - \hat{\theta}_{1,N}) 10^{(\hat{\theta}_{3,N} - x_i) \hat{\theta}_{4,N}}}{(g_{i,N}(x_i, \hat{\theta}_N))^2}$$

$$\frac{\partial y(x_i, \hat{\theta}_N)}{\partial \hat{\theta}_3} = - \frac{\ln(10) (\hat{\theta}_{2,N} - \hat{\theta}_{1,N}) (\hat{\theta}_{3,N} - x_i) 10^{(\hat{\theta}_{3,N} - x_i) \hat{\theta}_{4,N}}}{(g_{i,N}(x_i, \hat{\theta}_N))^2}$$

$$i = N_0 - 2, \dots, N + 1.$$

2: Solve the D-optimal problem (11) and find x_{N+1} .