Title

Automated deep brain stimulation programming with safety constraints for tremor suppression in patients with Parkinson's disease and essential tremor

Journal

Journal of Neural Engineering, 19(4)

ISSN

1741-2560

Authors

Sarikhani, Parisa
Ferleger, Benjamin
Mitchell, Kyle
et al.

Publication Date

2022-08-01

DOI

10.1088/1741-2552/ac86a2

Peer reviewed

HHS Public Access

Author manuscript

J Neural Eng. Author manuscript; available in PMC 2022 October 28.

Published in final edited form as:

J Neural Eng. ; 19(4): . doi:10.1088/1741-2552/ac86a2.

Automated deep brain stimulation programming with safety constraints for tremor suppression in patients with Parkinson's disease and essential tremor

Parisa Sarikhani[1], **Benjamin Ferleger**[2], **Kyle Mitchell**[3], **Jill Ostrem**[4], **Jeffrey Herron**[5], **Babak Mahmoudi**[1,6,8,*], **Svjetlana Miocinovic**[6,7,8]

[1]Department of Biomedical Informatics, Emory University, Atlanta, GA, United States of America

[2]Department of Neurosurgery, University of Pennsylvania, Philadelphia, PA, United States of America

[3]Department of Neurology, Duke University, Durham, NC, United States of America

[4]Department of Neurology, University of California San Francisco, San Francisco, CA, United States of America

[5]Department of Neurosurgery, University of Washington, Seattle, WA, United States of America

[6]Department of Biomedical Engineering, Emory University and Georgia Institute of Technology, Atlanta, GA, United States of America

[7]Department of Neurology, Emory University, Atlanta, GA, United States of America

[8]These authors contributed equally.

Abstract

Objective.—Deep brain stimulation (DBS) programming for movement disorders requires systematic fine tuning of stimulation parameters to ameliorate tremor and other symptoms while avoiding side effects. DBS programming can be a time-consuming process and requires clinical expertise to assess response to DBS to optimize therapy for each patient. In this study, we describe and evaluate an automated, closed-loop, and patient-specific framework for DBS programming that measures tremor using a smartwatch and automatically changes DBS parameters based on the recommendations from a closed-loop optimization algorithm thus eliminating the need for an expert clinician.

Approach.—Bayesian optimization which is a sample-efficient global optimization method was used as the core of this DBS programming framework to adaptively learn each patient's response to DBS and suggest the next best settings to be evaluated. Input from a clinician was used initially to define a maximum safe amplitude, but we also implemented 'safe Bayesian optimization' to automatically discover tolerable exploration boundaries.

[*]Author to whom any correspondence should be addressed., b.mahmoudi@emory.edu.

Conflict of interests

K M is a consultant on the innovation council for Boston Scientific and receives research support from Medtronic and Deep Brain Innovations. J O receives research support from Medtronic and Boston scientific, and is a consultant for Abbott. P S, B F, J H, B M and S M have no disclosures.

Main results.—We tested the system in 15 patients (nine with Parkinson's disease and six with essential tremor). Tremor suppression at best automated settings was statistically comparable to previously established clinical settings. The optimization algorithm converged after testing 15.1 ± 0.7 settings when maximum safe exploration boundaries were predefined, and 17.7 ± 4.9 when the algorithm itself determined safe exploration boundaries.

Significance.—We demonstrate that fully automated DBS programming framework for treatment of tremor is efficient and safe while providing outcomes comparable to that achieved by expert clinicians.

Keywords

closed-loop DBS; Bayesian optimization; intelligent systems; neuromodulation; wearable sensors

1. Introduction

Deep brain stimulation (DBS) surgery has become a standard treatment for neurological disorders such as Parkinson's disease (PD) and essential tremor (ET), to ameliorate tremor when medications are insufficient. DBS significantly improves both symptoms and quality of life, however to achieve therapeutic benefit, stimulation often requires time consuming programming by an expert [1]. DBS devices enable considerable customization of stimulation parameters including contact configuration (cathode and anode selections), current amplitude, pulse width and frequency allowing for customization of stimulation to account for variations in electrode placement, differences in local anatomy, symptom type, and severity [2]. Typically, a programming session involves a trial-and-error evaluation of therapeutic response (clinical benefit and unwanted side effects) at numerous stimulation settings. This is often performed over several sessions, which can be a challenge for patients who live far away from specialty care. In addition, evaluation of clinical response can be challenging given the subjective nature of visual observation to determine if tremor and other motor symptoms are responding to stimulation. Therefore, designing objective markers of therapeutic response to DBS is needed. Recently DBS device innovations (eight-contact electrodes, current fractionation, widened pulse width range and anodic stimulation) have significantly expanded the parameter space making programming even more complex [3]. These limitations suggest a need for an automated and patient-specific DBS programming framework that facilitates DBS programming without requiring an expert clinician.

We previously presented an automated DBS programming framework using an exhaustive grid search-based sampling strategy that mimics heuristic clinical DBS programming [4]. Although this automated framework was effective in programming DBS devices, sampling similar settings for all patients using a grid-search sampling method is a suboptimal approach since each patient responds to DBS differently [5]. Moreover, the number of required samples to converge to an optimal DBS setting was high, and we hypothesize that more advanced sampling and optimization techniques could improve the process. Two recent studies have assessed the efficacy of a closed-loop optimization algorithm for DBS programming using external motion sensor-based motor assessments in patients with PD [6, 7]. The details of the proprietary algorithm have not been published and the system required presence of a clinician to manually change the DBS settings based

on algorithm recommendations and if side effects occur. In addition, the algorithm-based DBS suggestions have only been tested with monopolar stimulation settings which may be suboptimal for some patients.

In this study, we combined the knowledge from clinical decision-making strategies with Bayesian optimization, to develop an automated real-time DBS programming framework that enables sample-efficient and patient-specific DBS programming to simultaneously ameliorate tremor and avoid side effects. Bayesian optimization has been successfully adopted for developing closed-loop neuromodulation applications in the context of optimizing the experimental design with closed-loop real-time functional magnetic resonance imaging (fMRI) [8], optimizing electrical stimulation for seizure control [9], searching through a large transcranial alternating current stimulation parameter space based on relative judgment [10], and for predicting optimal DBS parameters using fMRI data for PD patients [11]. The authors in [12] introduced a semi-automated approach to optimize DBS parameters in PD patients for minimizing rigidity and provided preliminary evidence on the utility of using Bayesian optimization in determining optimal DBS parameters. In addition, Bayesian preference learning was used in [13] for identifying personalized optimal stimulation patterns based on the participant's expressed preference for stimulation settings. The authors in [14] introduced a Bayesian adaptive dual control in a computational model of PD to reduce the beta power.

In addition, several recent studies explored DBS programming in closed-loop paradigm using tremor measurements. The authors in [15] investigated the use of isostable amplitude using computational models of ET patients to optimize DBS. Another study modeled the dynamics of patient tremor and their phase response curve to investigate the effect of phase-locked DBS in tremor suppression and proposed a closed-loop phase tracking stimulation regimens [16]. Several studies explored the utility of surface electromyography (EMG) and acceleration in tremor prediction and the design of a simple on-off DBS controller in closed-loop [17–21]. The authors in [22] used electrocorticography for sensing movement intention alongside with worn accelerometers and EMG sensors to deliver responsive closed-loop stimulation to treat tremor in a closed-loop fashion.

To the best of our knowledge Bayesian optimization with safety constraints has not been tested in the context of clinical DBS programming for tremor suppression. We hypothesized that implementation of a Bayesian optimization [23] algorithm for DBS programming for tremor suppression would have high efficiency (fewer samples than the grid search) and that safe programming (avoidance of uncomfortable side effects) can be achieved in an automated system using safe Bayesian optimization algorithm. We further provided clinical assessment of the closed-loop DBS programming framework in a cohort of 15 PD and ET patients.

2. Methods

2.1. Patient selection criteria and clinical experiment procedure

Patients with ET or tremor-dominant PD were recruited from a large academic movement disorders clinic. Patients had been implanted with Medtronic Activa neurostimulator systems

for at least 6 months and had DBS settings optimized during standard clinical programming visits prior to study enrollment. The Emory University Institutional Review Board approved the study and all patients signed written informed consent.

At the beginning of each experimental session, DBS was turned off, and clinical setting stimulation effects washed out for 10 min. Patients were asked to hold their tremor or PD medications for at least 12 h prior to testing to avoid medication fluctuations during optimization. DBS optimization was performed in one lead for each patient, contralateral to the arm with more severe tremor. The non-tested lead was kept off during optimization unless rest tremor was too bothersome to sit comfortably. DBS implantable pulse generator (IPG) was reprogrammed to create four groups (four contact configurations with amplitude control) that the optimization algorithm could explore (at the beginning of the experiment, the four groups are set as monopolar contact configuration, where the IPG case is set as an anode and each of the four contacts are set as a cathode, and if 'advanced' stimulation was needed based on algorithm's decision scheme, we set the four groups as bipolar or multipolar contact configurations). Stimulation pulse width and frequency were not changed during optimization and were the same as the patient's clinical setting. For phase I experiments, clinician determined maximum allowable amplitude for each stimulation group that could be safely sampled by the automated optimization algorithm (to prevent the automated system from inducing severe side effects). In phase II experiments, the maximum allowable amplitude was set to 5 V for all groups, and the safe Bayesian optimization algorithm was utilized to avoid inducing severe side effects (additional safety feature allowed rapid stimulation shut off if necessary).

During optimization, two standard clinical motor tasks were performed at each stimulation setting to assess tremor, depending on each patient's tremor profile (rest, arms extended, arms flexed, or finger-nose motion). A commercial smartwatch (LG-W100) worn on patient's wrist was used to determine a tremor score using a previously validated classifier [4] and this score was used as input into the optimization algorithm (figure 1(a)). A clinician blinded to the stimulation setting also scored the tremor during optimization using Fahn–Tolosa–Marín (FTM) rating scale [24] to further assess previously validated tremor classifier [4]. At each stimulation setting, the patient reported stimulation-induced side effects (typically tingling or muscle contractions in the face, arm or leg) and rated them on a scale from 0 to 3 (none, transient or mild, moderate, severe). At the end of the optimization session, DBS IPG was set to the best 'automated setting' and a clinical tremor exam and objective watch tremor measurement were performed after a 5 min wash-in (examiner and patient were aware of the stimulation condition). Clinical tremor exam for both PD and ET patients was a subset of FTMscale and included rest, postural, and action arm and leg tremor contralateral to DBS lead, handwriting (if dominant hand tested) and spiral and line drawings. DBS IPG was then set to patient's 'clinical setting' and after another 5 min wash-in period, tremor was reassessed by exam and watch (for the first two patients, tremor assessment at clinical setting was done at the beginning of the visit, but protocol was changed for subsequent subjects to facilitate direct comparison between automated and clinical settings).

2.2. Automated DBS programming framework: software design

We performed these automated DBS programming experiments using a custom software application developed for a Windows PC (figure 1(b)). This application collected patient-reported side effect data and inertial measurement unit (IMU) data from the smartwatch. Side effects were entered through a user interface and constituted a total of 13 common acute side effects of DBS therapy and included reports of magnitude (‘0: none’, ‘1: mild’, ‘2: moderate’, ‘3: severe’), type (‘paraesthesia’, ‘muscle spasm’, ‘speech’, ‘vision’, ‘dizziness’, ‘dyskinesia’), and body location (‘head’, ‘arm’, ‘leg’, ‘torso’). IMU data, constituting three dimensions each of accelerometer and gyroscope data for a total of six channels, was streamed to the study PC via Bluetooth at 100 Hz for processing and feature extraction. All data was logged on receipt in comma-separated values (CSV) format, as were DBS stimulation parameters.

The software application made use of a previously developed C# application programming interface [25–27] for interfacing with the Nexus-D, a Medtronic research communication bridge that allows an application to update IPG stimulation parameters. IMU data collection and interfacing with the DBS device through the Nexus-D was conducted through this application. Time-series IMU data review and side effect inputs were conducted using an application written in Python to take advantage of superior data processing and capacity to deploy advanced machine learning and optimization techniques. Lab streaming layer, a publicly available library for cross-platform port handling and communication of time-series data in research applications, formed the interface between these applications.

Using this combination of side effect data and tremor severity estimate, a quantitative therapeutic value was derived for each DBS setting using Python. The details of this derivation are described in more detail in the following sections. Updated settings were forwarded to the C# application. Both the therapeutic value of the preceding DBS settings and the recommended next settings were recorded at this point. Following a brief final review and safety check, these settings were then communicated via USB connection to the Nexus-D. The C# application was also capable of manual override in case of emergency.

2.3. GPR modeling of the effect of DBS settings using a quantified objective measure

Gaussian process regression [28] is a nonparametric, Bayesian regression approach, which is well-suited for small datasets and provides the measurement uncertainty for the predictions. In this study, the patient-specific GPR models used the Matern kernel function [23] and were trained using the cumulatively collected samples $D = \{(x_i, y_i) \mid i = 1, \ldots, n\}$ from each patient, where x_i were stimulation parameters, i.e. stimulation amplitude and stimulation contact configuration and y_i represented the corresponding combined objective measure as defined in equation (4). A gaussian process (GP) is a nonparametric model that is fully characterized by its mean and covariance function as following

$$f(\boldsymbol{x}) \sim \mathrm{GP}(m(\boldsymbol{x}), K(\boldsymbol{x}, \boldsymbol{x}')), \tag{1}$$

where we can define the mean function as $m(x) = E[f(x)]$ and the covariance function as $K(x, x') = E[(f(x) - m(x))(f(x') - m(x'))]$. Here, we used Matern kernel function as in

$$K_{\text{MATERN }3}(\boldsymbol{x}, \boldsymbol{x}') = \sigma_f^2 \exp(-\sqrt{3}r(1+\sqrt{3}r)) + \sigma_n^2 I, \tag{2}$$

where $r^2 = (\boldsymbol{x} - \boldsymbol{x}')^{\mathrm{T}} \Lambda (\boldsymbol{x} - \boldsymbol{x}')$ and Λ is the diagonal matrix of squared length scales. The output variance σ_f^2, the length-scales, and the noise variance σ_n^2 are hyperparameters of the covariance function. By incorporating the knowledge from the training data (prior distribution in equation (1)), we can make predictions at any new test point (x_*, f_*), where $f_* = f(x_*)$. The predictive conditional distribution of f_* given the training data and test input is calculated as in

$$f_* \mid X, y, X_* \sim N\left(\overline{f_*}, \operatorname{cov}(f_*)\right), \tag{3}$$

where $\overline{f_*} = E[f_* \mid X, y, X_*] = K(X_*, X)\left[K(X, X) + \sigma_n^2 I\right]^{-1} y$,

$\operatorname{cov}(f_*) = K(X_*, X_*) - K(X_*, X)\left[K(X, X) + \sigma_n^2 I\right]^{-1} K(X, X_*)$, and σ_n^2 denotes the noise variance.

This GPR modeling technique is used as a surrogate model for Bayesian optimization described in the following section. We employed the GPflow library [29] for implementing GPR models. We defined a combined quantitative measure of clinical efficacy consisting of a quantitative objective tremor score, measured by the smartwatch IMU and patient-reported side-effects, with the goal of maximizing tremor improvement and minimizing side-effects. The quantification measure of DBS setting value, J_{DBS_i}, is calculated based on the results of the tremor assessment tests while the patient's IPG was active in a particular DBS setting DBS_i as in

$$J_{\text{DBS}_i} = J_{\text{tremor}_i} + J_{\text{SE}}. \tag{4}$$

Each DBS setting DBS_i is evaluated based on the tremor score improvement, J_{tremor_i}, which is a baseline subtracted tremor severity score as in equation (5), and J_{SE} which is the patient-reported side effect severity scores defined as in equation (6).

A predictive model of clinical tremor assessment from IMU data was trained and validated in a previous study [4], where features from accelerometer and gyroscope data were used to train an ordinal multinomial logistic regression classifier based on the neurologist's provided tremor ratings [4]. To directly compare the performance of our Bayesian automated DBS optimization framework with the state-of-the-art automated DBS programming framework introduced in [4], which uses a grid-based search approach, we used the same classifier [4] in order to avoid introducing new parameters to the system. The output of this classifier was used to calculate the tremor score improvement, J_{tremor_i}, as in equation (5).

The term $\left(\text{tremor}_{\text{DBS}_i}\right)$ in equation (5) is the average watch tremor severity score (predicted from the classifier) over the selected tasks while the patient's IPG was active with the ith DBS setting DBS_i. The term tremor_0 is the average watch tremor severity scores (predicted

from the classifier) over all selected tasks with inactive IPG that reflects patients' baseline tremor score obtained at the beginning of optimization session:

$$J_{\text{tremor}_i} = \text{tremor}_{\text{DBS}_i} - \text{tremor}_0 . \tag{5}$$

The magnitude of the overall baseline subtracted tremor score $\left(J_{\text{tremor}_i}\right)$ shows the level of change in the average tremor score comparing to the baseline and a negative sign reflects tremor improvement compared to the baseline. The lower the J_{tremor_i}, the more clinical benefit the DBS_i setting provides.

In addition, each DBS setting DBS_i is penalized by the patient-reported side effect severity scores. We defined the term J_{SE} in equation (4) as follows based on patients' reports:

$$J_{\text{SE}} = \begin{cases} 0 & \text{if no SE} \\ 1 & \text{if mild SE} \\ 4 & \text{if moderate SE} \\ \text{inf}(5 \text{ in practice }) & \text{if severe SE.} \end{cases} \tag{6}$$

Watch tremor severity scores are on a scale of 0–4, so the J_{tremor_i} term could get any value in the [−4, 4] interval depending on the baseline score and the tremor score in the DBS_i setting. Specifically, $J_{\text{tremor}_i} = 0$ means no tremor improvement, $J_{\text{tremor}_i} < 0$ reflects a tremor score improvements compared to baseline, and $J_{\text{tremor}_i} > 0$ corresponds to cases where the watch tremor score is worse than the watch tremor score at baseline. If the patient experiences some level of side effect, we penalize J_{tremor_i} by adding a positive value to it. If the side effect is mild, we penalize it by 1, meaning that a DBS setting with a score of 1 for tremor improvement with mild side effect will have a total score 0 which is similar to a setting with no improvement and no side effects. If the side effect is moderate, we penalized it by 4 because any amount of tremor improvement with moderate SE will be considered as untenable for clinical use (resulting in a non-negative J_{DBS_i} score). If the SE is severe, we penalize it even more to prevent the optimizer from testing that area again.

The quantified objective measure in equation (4) was calculated for each DBS setting tested on the patients and the cumulatively collected samples were used to train a GPR model that models patients' response to DBS. The mean surfaces of GPR models of the combined objective measure defined in equation (4) capture both the effect of baseline-subtracted watch tremor score and the side effect severity scores simultaneously and justifies the use of the combined objective measure for Bayesian optimization to ameliorate tremor while avoiding side effects (figure 2). Furthermore, the mean surface of the GPR model varies across patients (figure 2). Specifically, the general shape of the surface, the location of the optima, and the maximum tolerable boundaries vary between individuals reflecting their unique response to stimulation profile. These variations are partly due to disease type and tremor severity, DBS lead position (which varies even for patients with the same target nucleus), and individual anatomy. This subject variability necessitates designing a

patient-specific DBS optimization framework with an adaptive sampling strategy, while still remaining sample-efficient. Due to variability in patients' responses, a grid search-based approach with 1 V amplitude increments for each contact configuration that was utilized in a previous similar study [4] was hypothesized to be inefficient (figure 2(a)). The Bayesian optimization which we have utilized in this study evaluates more samples in areas with greater chance of tremor improvement and searches for the optimal point with a finer resolution of 0.2 V increments in amplitude for each contact configuration (figure 2(b)).

2.4. Bayesian optimization

We formulated the automated DBS programming as a global optimization problem over the stimulation parameter space, D, as in:

$$\min_{x \in D} f(x), \tag{7}$$

where $f(x)$ was the objective measure that represented the desired clinical outcome, and D is the two-dimensional space of the DBS parameters including stimulation amplitudes and contact configurations. One of the main challenges in solving this problem was that the functional relationship between the DBS parameter space and clinical outcome was not known. Bayesian optimization is a non-parametric global optimization approach that is suitable for optimizing black-box objective functions that are unknown or expensive to evaluate.

The objective function $f(x)$, which represents a mapping between the DBS parameters and the clinical outcome, is unknown and does not have a closed form. Although $f(x)$ is unknown at every $x \in D$, we can observe its measurements at sampled DBS settings (the objective measure as described in equation (4) is a measurement of $f(x)$ at the suggested DBS settings DBS_i by the optimizer). Bayesian optimization proceeds by maintaining a probabilistic belief over $f(x)$ by building a GPR surrogate model as described in section 2.3 using the cumulatively collected data from each patient. The GPR model prescribe a prior belief over the possible objective functions given the cumulatively collected data. In a previous study we characterized the functional relationship between the clinical outcome and DBS parameters using the GPR modeling approach [5].

The Bayesian optimization algorithm is based on a sequential decision-making process to search for the optimal stimulation parameters in two steps. First, it builds a surrogate probabilistic model of the latent objective function $f(x)$ based on the available data at each iteration and sequentially retrain the model as more data is observed. Second, it proposes the next DBS setting to be evaluated by optimizing a surrogate-dependent acquisition function. The acquisition function assesses the utility and the informativeness of the candidate points for the next evaluation of the objective measure ($f(x)$) by leveraging the uncertainty in the posterior to guide exploration [23].

During the burn-in phase of Bayesian optimization, the objective function was evaluated at predefined stimulation settings in a randomized order (at 40% and 80% of maximum amplitude for each contact configuration). Then, a GPR prior based on these initial evaluations of the burn-in phase was employed. Thereafter, a new stimulation setting was

sequentially selected by optimizing the acquisition function to be evaluated at next iteration. At each iteration, we augmented the dataset, updated the surrogate GPR prior, and optimized the updated surrogate-dependent acquisition function to suggest the next samples to be tested in the patient until convergence. Our stopping criteria is explained in section 2.6. For the first round of experiments, with amplitude limits determined by a clinician, we used the expected improvement (EI) acquisition function [30] which automatically balanced exploration versus exploitation. The EI acquisition function calculates the expectation of improvement over the current best observation with respect to the predictive distribution of the surrogate model and is defined as in:

$$\begin{aligned} \mathrm{EI}(x) &= E[\max(f_{\min} - m(x), 0)] \\ &= (f_{\min} - m(x))\Phi(z) + \sigma(x)\phi(z), \end{aligned} \tag{8}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the standard normal density and distribution functions, respectively. In equation (8), $z = \frac{f_{\min} - m(x)}{\sigma(x)}$, $m(x)$ is the predictive mean and $\sigma(x)$ is the predictive standard deviation of a point $x \in D$ and $f_{\min}$ is the optimum observed value. We implemented the EI acquisition function using the GPflowOpt library [31].

The global optimization problem defined in equation (7) is straightforward to solve using Bayesian optimization algorithm if the parameter space is fully defined. To show the feasibility of Bayesian optimization as the core of the automated DBS optimization framework during the first phase of the experiments, the maximum amplitude for each contact configuration is defined at the beginning of the experiment by the expert neurologist and will stay fixed during the experiment as shown in red dashed boundaries in figure 3. The minimum exploration boundary of stimulation amplitude in the parameter space is set to 0.5 V (that is the dashed horizontal red line at 0.5 V in figure 3); meaning that the optimizer will not explore the effect of DBS settings with amplitudes smaller than 0.5 V. With more samples being collected at each iteration, the underlying GPR model gets updated. At each iteration, the updated GPR model is used to build the acquisition function and get the next DBS setting suggestions to be evaluated at the next iteration. DBS settings are adaptively sampled during the DBS optimization session based on patient's response in a patient-specific manner (figure 3).

Since the objective function $f(x)$ is unknown, our Bayesian optimization algorithm does not assume convexity. As a result, if multiple optima are found, the setting with the lowest amplitude was selected as the optimum automated setting to ensure lower power is used when possible.

2.5. Safe Bayesian optimization

The DBS optimization problem is safety-critical, as there is a safety constraint for each DBS setting. The safety constraints are defined by the side effects that patients may experience for each DBS setting. In the second phase of the experiments, instead of using the clinician-defined safe exploration boundaries, we modified the problem statement as a global optimization problem with safety constraint as in:

$$\min_{x \in D} f(x); \text{ Such that } g_i(x) < 3 \text{ for } i = 1, 2, 3, 4, \tag{9}$$

where $g_i(x)$ is the magnitude of patient-reported side effect for each DBS setting x and i is the contact configuration number. Since the safety boundaries for each contact configuration are different, a separate GP model was trained for each contact configuration based on the patient-reported side-effects. To solve the above constraint optimization problem, we used safe Bayesian optimization [32, 33] which is an extension of regular Bayesian optimization. We used the GPflow [29] and GPflowOpt [31] libraries for our implementation of the algorithm.

Safe Bayesian optimization combines a GP model of the safety constraints with discretization of the parameter space to define a set of DBS parameters S_n, called the safe set, with a high probability to satisfy the safety constraints [32]. The safe set was defined by the upper bound of the safety GP models ($g_i(x)$) and contained the points where the GP upper bound was smaller than the safety threshold. Our parameter space for the DBS programming was discrete, with four contact configurations and stimulation amplitudes with 0.2 V increments. The safe Bayesian optimization algorithm defined two sets of parameters within the safe set called potential minimizers and expanders. The set $M_n \subseteq S_n$ contains potential minimizers that is the parameters that could potentially obtain the minimum within the current safe set. The minimizers set (M_n) was defined by the mean and confidence interval of the GPR model of the objective measure, $f(x)$, and contained the subset of safe parameters in which the lower confidence bound of the GPR model was lower than the upper confidence bound at the best measurement at each iteration. The expander set $G_n \subseteq S_n$ is considered to be the points that if tested, their measurements would lead to values in the lower confidence bound and hence potentially expand the safe set [33]. In this paper, we directly used the confidence bounds of the GPR models to define the aforementioned sets, which is mainly used in practice with limited prior knowledge of the underlying objective and safety models (considers the Lipschitz constant to be infinity). This modified version has been shown to be more aggressive in expanding the safe set [33]. Hence, to further ensure safety, we considered the maximum allowable amplitude expansions based on the maximum severity of the patient-reported side effects for each contact configuration. The safe set would only expand by the minimum of the safe set suggested by the safe Bayesian optimization algorithm and the maximum allowable constant defined as follows. In general, the lesser the reported side effect, the more the parameter space could expand. In other words, the expanded space gets closer to the safety boundaries as the patient begins to experience mild side effects by increasing the stimulation amplitude. Hence, the expansion should be done with more caution. Another consideration in selection of the constant was that patients were more likely to experience more severe side effects during the monopolar stimulation settings than advanced stimulation settings. Therefore, the constant during the monopolar stimulation was selected to be smaller. If the patient experienced no side effects for a particular contact configuration, then the constant was set to 1 V for monopolar stimulation and 1.5 V for the advanced stimulation settings, respectively; these amplitudes represent the maximum allowed jump in the stimulation amplitude for each contact configuration. If the maximum side effect severity level was

mild, then the constant was set to 0.5 V for monopolar stimulation and 0 V for advanced stimulation settings (zero because we aimed for no side effects at optimal stimulation setting during advanced stimulation; during monopolar stimulation mild side effects were tolerated during optimization as results were informative for selection of advanced stimulation configurations). If the patient experienced moderate or severe side effects either during the monopolar or advanced stimulation, the constant was 0 V and the expansion of the stimulation amplitude was stopped for the corresponding contact configuration.

Safe Bayesian optimization starts with evaluating a set of initial parameters that is known to be safe, which defines the initial safe parameter space. Here, we started with testing 1 V for each of the four contact configurations as the initial safe set of parameters (S_n). The initial evaluations were used to train GP models on the safety constraints which were used to safely expand the parameter space at each iteration. After the safe expansion of the parameter space, the next DBS settings to be tested were suggested by optimizing the surrogate-dependent acquisition function at each iteration. As mentioned above, we used GPR modeling technique to model the constraint/safety functions. A common assumption in training GPR models is that a GP prior is zero-mean. However, this assumption did not apply to the DBS programming problem since the side effect severity report was a monotonically increasing function of stimulation amplitude. Hence, we fitted a second-degree polynomial function of the collected samples and used that as the prior mean of the safety GP models. In addition, in phase II of the experiments, we changed the acquisition function to minvalue entropy search [34] as it was more sample efficient and had a more exploratory behavior that is required given the nature of safe Bayesian optimization algorithm with gradual expansion of the parameter space.

A visual representation of the sampling behavior and safe exploration boundary expansion of the automated framework is presented for some sample iterations in figure 4 for patient 14. Each figure shows the mean surface of the GPR model that gets updated as we collect more data at each iteration. Note that the safe boundaries shown in dashed red line are updated after evaluating the suggested DBS setting at each iteration.

In both phases of the experiments, we considered some modifications of the regular (safe) Bayesian optimization algorithm to account for the requirements of the automated DBS programming in practice. First, we considered a discrete parameter space including four contact configurations and stimulation amplitude with 0.2 V. The optimization of the acquisition function is performed by evaluating the acquisition function at every setting in the parameter space at each iteration and the next best sample is selected using the selection of the best strategy. One challenge in designing Bayesian optimization in discrete parameter spaces is the suggestion of repeated samples. In order to avoid suggesting repeated samples, we used a rank and select strategy; that is if the best sample suggested by optimizing the acquisition function is already sampled, the next best sample will be suggested.

2.6. Stopping criteria and advanced optimization

Our automated programming framework combined prior knowledge from standard clinical DBS programming approaches with Bayesian optimization. Clinical DBS programming is guided by clinical decision-making strategies to maximize stimulation benefit and minimize

side effects [35]. During clinical DBS programming sessions, the clinician often performs a monopolar screening, where each of the four electrode contacts are set as cathode and IPG case as anode. Clinicians evaluate the tremor suppression benefits and side effects by incrementally increasing the amplitude for each of the four monopolar contact configurations. If a monopolar configuration leads to a satisfactory tremor suppression without side effects, this setting is chosen for chronic stimulation. If not, contact configuration may be changed to one of the advanced optimization settings (bipolar, double-monopolar, and double-bipolar) based on patients' responses to monopolar stimulation (figure 5).

Similar to the standard clinical approaches, the proposed automated DBS programming framework started with the monopolar screening with four different contact configurations. The monopolar optimization was terminated either once the stopping criteria was satisfied or the predefined budget of maximum 30 iteration was exhausted (to avoid patient fatigue). The stopping criterion was defined as no objective score improvement of 0.3 or greater for five consecutive iterations. The threshold of 0.3 was defined by our expert movement disorder neurologist that reflected a clinically meaningful score improvement.

After completing the monopolar stimulation trials, the advanced stimulation module used the data that was collected during the monopolar stimulation to determine if advanced stimulation was necessary, and which contact configurations should be utilized (figure 5). If there was at least one DBS setting with sufficient therapeutic effect and without any side effect (acceptable monopolar setting), then the monopolar setting with the lowest average tremor score was selected as the optimized setting. Otherwise, if there was sufficient therapeutic effect with the presence of side effects, then bipolar stimulation was suggested. The sufficient therapeutic effect was dependent on the baseline score. If the average baseline tremor score was less than 1, then sufficient therapeutic effect was defined as 50% improvement in watch tremor score. Otherwise, having a watch tremor score of less than 1 for all of the tremor assessment task was considered as having sufficient therapeutic effect. In cases where there was no setting with sufficient therapeutic effect and no side effect for amplitudes less than 4 V, then double-monopolar stimulation was suggested. If there was no setting with sufficient therapeutic effect and side effects at less than 4 V were present, double-bipolar stimulation was suggested. If no acceptable monopolar setting was identified, the advanced stimulation suggestion script (darker gray area in figure 5) provided four advanced contact configurations based on clinical heuristics. Since the number of possible advanced stimulation settings is high and due to the limitations of the Medtronic research communication bridge, we could only test four contact configurations at each round of automated DBS programming. We used the advanced stimulation suggestion script to use the clinical heuristics and narrow down the number of advanced settings to 4 best settings to be tested (as depicted in the blue box in figure 5). Testing new settings continued until either the stopping criteria was satisfied, or the maximum number of iterations was reached.

3. Results

We recruited 15 patients (nine with tremor-dominant PD and six with ET) with the average age of 70 ± 9 years (range 57 − 85) to undergo the automated DBS optimization. All ET

patients and one PD patient had leads implanted in ventral intermediate nucleus of thalamus (VIM), while the remaining PD patients had leads implanted in subthalamic nucleus (STN). The average time since DBS lead implantation was 52 ± 19 months (range 30–105 months). We evaluated the performance of the automated DBS optimization framework in two phases. In the first phase, the software used the clinician-defined maximum tolerable amplitudes for each contact configuration to define the safe boundaries of the parameter space. The main goal of the first phase of the experiments was to evaluate the performance of the automated framework in finding the optimized settings using Bayesian optimization algorithm. Data from ten patients (five PD and five ET) were acquired for the first phase (table 1). We further expanded the work and used safe Bayesian optimization to gradually expand the parameter space and automatically discover a safe and tolerable parameter thus avoiding severe side effects as reported by the patient. Seven patients (five PD and two ET; two from phase 1) underwent the automated DBS optimization in the second phase of experiments (table 2).

3.1. Quantifying tremor response to stimulation

The automated DBS optimization framework automatically quantifies and calculates the target objective measure that includes tremor scores and side effects (figure 1(a)). We confirmed that the tremor classifier [4] performed well in this cohort of patients and its estimated tremor scores matched well with clinician scores ($r^2 = 0.69$; figure 6). The clinician tremor score includes only tremor assessment tasks during the automated programming sessions and was used to further validate watch tremor classifier. The clinician administered FTM tremor scale is a more comprehensive examination consisting of a subset of FTM scale items used to evaluate tremor severity before and after automated programming session. Only the automated classifier tremor scores were used as the input into the optimization algorithm.

3.2. Comparison of the clinical settings and the automated settings

There was a statistically significant improvement in tremor scores from baseline (no stimulation) to the best automated setting, using both the objective watch scores and blinded clinician scores during both phases of the experiment (figure 7) (the clinician tremor score is the score for selected tremor assessment tasks during the experiment where the clinician was blinded to the DBS settings). The patients also underwent a comprehensive tremor assessment exam at baseline (no stimulation), best automated setting, and their chronic clinical settings (tables 1, 2 and figure 8). We demonstrate that best automated setting and clinical setting significantly reduce the tremor to the same extent (in other words, residual tremor at automated setting was comparable to tremor at clinical setting) (figure 8).

In phase I experiments with the clinician-defined safe and tolerable exploration boundaries, two patients preferred the automated setting, five had no preference, and three preferred their clinical settings. In phase II experiments with automated discovery of the safe exploration boundaries, three patients preferred the automated setting, three had no preference, and one preferred the clinical setting.

3.3. Speed of convergence of the automated DBS programming system

We hypothesized that our Bayesian DBS programming framework would improve sample efficiency compared to the grid search-based method in terms of the number of stimulation settings had to be tested to arrive at the optimal solution (not in terms of the required time). The grid-search approach closely resembled clinical monopolar mapping (testing all contacts and amplitudes, 0–5 V, in 1 V increments). We could not compare the current algorithm directly against clinical monopolar mapping in terms of time since these patients had already been clinically optimized. In order to provide a fair comparison between the two approaches, we only considered the number of required samples during the monopolar programming. Grid search algorithm tested in a prior study (in a different cohort of patients) required 25.2 ± 4.8 samples on average [4], while Bayesian automated programming in this study used 15.1 ± 0.7 (phase I), and 17.7 ± 4.9 samples (phase II).

4. Discussion

In this pilot study, we describe and evaluate an automated and patient-specific DBS programming framework for tremor treatment in 15 patients with PD or ET. A fully automated system with the Nexus-D communication bridge was developed that automatically activates the patients' IPG with the optimizer recommended DBS settings. We showed that DBS programming framework using Bayesian optimization was able to find DBS settings that were comparable in efficacy to clinical settings (previously determined by expert clinician programmers). Bayesian optimization was more efficient than previously tested grid-search method. We also describe how to use safe Bayesian optimization to automatically find safe stimulation boundaries. Finally, by incorporating the information from monopolar stimulation and clinical heuristics, we were able to add advanced DBS contact configurations (bipolar, double monopolar) that some patients require for optimal therapy into the automated DBS programming workflow and perform further optimization using four advanced DBS contact configurations. These developments may reduce the need for an expert clinician programmer to be present at the DBS programming session to perform DBS device control, symptom and side effect assessment, DBS programming decision making, and defining the safe and tolerable amplitudes for each contact configuration.

A physician can manually explore any number of settings; they are limited by the time available for a clinical visit and the patient's ability to actively participate. The purpose of the algorithm was to test only the settings most likely to yield the optimal solution. Because of the type of DBS device that the patients were implanted, the algorithm was limited to amplitude changes in four contact configurations during each optimization, so we tested four configurations during monopolar stimulation, and additional four during advanced stimulation if monopolar did not yield the optimal setting. Future DBS devices may provide more flexible interfaces for automated stimulation adjustments allowing wider parameter exploration.

Bayesian optimization has unique properties that make it a suitable choice to be employed at the core of an automated DBS optimization framework. Bayesian optimization is a sample-efficient and global optimization algorithm that is suitable for cases where the objective

function is unknown or expensive to evaluate (the patients' response to DBS settings are unknown prior to testing and evaluating patients' responses to DBS settings is expensive from optimization standpoint as prolonged sessions are fatiguing which may compromise the accuracy with which tremor is assessed during testing). We confirmed our hypothesis that Bayesian optimization was more sample-efficient than the state-of-the-art grid-search sampling strategy introduced in [4] which closely resembled clinical monopolar mapping and directly compared the results in terms of the number of required samples to be tested to arrive at the optimal solution. Other recent studies investigated the utility of developing an objective measure for the automated selection of DBS parameters [36] and introduced a computer-guided DBS programming framework that is designed based on the clinical DBS programming strategies for the monopolar survey [37] using a grid-search approach resembling the standard clinical approaches. Their sampling strategy was a grid-search approach with 0.5 V and 0.3 V amplitude increments, respectively, leading to an even larger number of required samples to be tested compared to the grid-search approach with 1 V increment introduced in [4]. Two recent clinical papers compared their proprietary algorithms with the standard of care (SoC) DBS programming in terms of the number of steps (stimulation settings) required to be tested to arrive at an optimal solution [6, 7]. The SoC was designed to be similar to the grid-search based approach. However, we could not conduct a direct and fair comparison with approaches used in [6, 7] since their parameter space was different (eight monopolar contact configurations and stimulation current (mA)). Due to very different workflows in SoC and the proposed closed-loop algorithm, the authors did not perform significance testing between the two programming modalities for time consumption. Here, we could not compare the current algorithm directly against clinical monopolar mapping in terms of time since these patients had already been clinically optimized. Another recent work [38] used Bayesian optimization to develop a semi-automated approach for optimizing DBS parameters and provided preliminary data that shows the efficacy of Bayesian optimization in DBS programming. Here, we presented and evaluated the utility of Bayesian optimization in a fully automated DBS programming framework for tremor in a cohort of 15 PD and ET patients.

We further showed that employing safe Bayesian optimization algorithm enables unsupervised determination of safe stimulation parameters. Safe Bayesian optimization is less sample efficient in nature than the regular Bayesian optimization as it gradually expands the parameter space. To improve the sample efficiency, we used ideas from [33] to balance the tradeoff between exploring, expanding, and optimizing in addition to using a more efficient acquisition function (min-value entropy search). We showed that although safe Bayesian optimization in phase II experiments required more samples to converge than the regular Bayesian optimization in phase I, it is still more sample-efficient than the grid-search approach.

Incorporating clinical heuristic into the optimization pipeline allowed us to efficiently explore advanced contact configurations (bipolar, double monopolar). The number of possible contact configurations beyond simple monopolar is very large and we could only test a relatively small number given that patients fatigue after prolonged and repetitive testing. As a result, we used information obtained from monopolar stimulation to determine which contact configurations should be tested during advanced stimulation using clinical

guidelines programmed into the programming platform rather than allowing the optimizer to make this decision. Integrating other algorithms that are more efficient for high-dimensional parameter spaces or other methods such as image-guided programming [38] to more efficiently reduce the dimensionality of the parameter space before performing the DBS optimization is a promising approach for future applications.

Our DBS programming framework relied on automated tremor detection using a wrist-worn sensor which has challenges [39], but we demonstrate that some of the issues can be overcome through GPR modeling. Tremor assessment tasks need to be synchronized with IMU recordings during the assigned tasks so patients need to be instructed to start and stop at appropriate times. Furthermore, artifact and voluntary movements unrelated to assigned tasks are often included in raw IMU data which affects the prediction of watch tremor scores, while an expert clinician programmer could detect those unrelated movements and ignore them while making judgment about the tremor severity scores. In this study, we monitored the patients and gave instructions to minimize the unrelated movements (including repeating a task if performed incorrectly), however variability in task performance (e.g. speed of movement) particularly during kinetic tremor assessment led to score predictions that were at times inconsistent with clinical scoring. Another challenge with automated tremor detection is that tremor severity can change depending on patient's internal state; for example, there can be less tremor when relaxed, and more tremor when nervous or talking, regardless of DBS settings. For example, tremor intensity during the optimization session varied significantly in patients 03 and 09 regardless of applied stimulation. A clinician can easily incorporate this information into clinical decision-making however an external sensor is agnostic to patient's internal state. To compensate for imperfections with tremor scoring, we employed the GPR model as the surrogate model of Bayesian optimization which is robust to noise of observations. Our results confirmed that the automated DBS programming method could identify effective DBS settings even in the presence of the measurement and prediction (i.e. classifier) noise. GPR model takes the uncertainty of observations into account and the model can be trained in a patient-specific manner which makes it suitable for the DBS optimization application.

The GPR modeling could be integrated in clinical decision-making process as a visualization technique that provides insight into the patient's response to DBS even without utilizing the fully automated platform. This would be particularly useful when addressing symptoms other than tremor which are even harder to quantify using sensors (e.g. bradykinesia or rigidity in PD). This visualization technique could provide insight into the spatial information (location of the active contacts) for clinicians that may not be straightforward to capture using the traditional clinical programming approaches. There have been attempts to improve visualization of DBS programming outcomes for clinicians especially with more complex segmented electrodes (e.g. [40]). We propose that GPR models could be used to not only track clinical responses but also provide suggestions for the clinicians for further DBS parameter exploration.

The success of the optimization algorithm will also depend on the choice of tremor assessment tests which are performed at each stimulation setting. In this study, we used two out of four available tremor tests (rest, postural extended, postural flexed, and kinetic)

that were integrated into the automated programming software system, based on patient's clinical presentation. Although the limited set of tremor assessment tests was sufficient to evaluate DBS response in majority of patients, some patients may require other types of tests including spiral and line drawing or handwriting tests to effectively find an optimal DBS setting. For example, patients 04 and 07 had worse tremor control on automated setting than on clinical setting, likely because they had more tremor on handwriting and spiral drawing tasks, which were not tested during automated programming session.

The automated system relied on patient reports of side effect severity as part of the combined objective measure for Bayesian optimization. Although this approach has been effective in avoiding parameters that lead to side effects in our study, this is a potential limitation since some patients find it difficult to give a score to the side effects that they experience. Developing automated side effect detection techniques could be possible for certain types of side effects (e.g. muscle contractions measured by EMG [41] or stimulation outside DBS target volume estimated by computational DBS activation models [38]), and could further streamline implementation of automated programming.

The objective measure defined in this study is based on aggregation of two terms including the baseline subtracted tremor score and patient-reported side effect severity score. This works since the two aggregated terms are in a comparable range. If another out of range term needs to be added to the objective measure, methods like adding a multiplier should be used to map the new term to the same range. Moreover, the work in this study can be extended to multi-objective optimization to incorporate more advanced objective measures of the clinical outcomes.

Finally, the proposed automated DBS programming framework could be beneficial for remote DBS programming for patients with limited access to the clinic. For example, a smartwatch could be mailed to the patient prior to a remote programming session or patient's own phone could provide accelerometer signal to quantify tremor. The optimization algorithm could be implemented as a standalone system providing guidance to the remote programmer, or even incorporated into the remote programming software.

5. Conclusion

This study developed and tested automated and patient-specific closed-loop DBS programming framework based on Bayesian optimization. This approach was more efficient than grid search method employed in clinical practice, and it yielded comparable clinical outcomes for tremor reduction as traditional clinical programming. Using such system would eliminate the need for an expert clinician programmer to be present at the DBS programming sessions. This would be particularly valuable for patients without easy access to DBS center such as those living in remote geographical locations or patients receiving care via telemedicine. Automated DBS programming methodologies will be of increasing importance as next generation DBS systems expand the number of possible parameters for delivering precise, optimized therapy to patients.

Acknowledgments

This work was in part supported by the NIH Award Grants 1P50NS098685 and R01EB028350. We thank the patients for participating in the study. We thank Medtronic for providing Nexus adapter and technical assistance.

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

References

[1]. Hunka K, Suchowersky O, Wood S, Derwent L and Kiss ZHT 2005 Nursing time to program and assess deep brain stimulators in movement disorder patients J. Neurosci. Nurs. 37 204–10 [PubMed: 16206546]

[2]. Picillo M, Lozano AM, Kou N, Munhoz RP and Fasano A 2016 Programming deep brain stimulation for tremor and dystonia: the Toronto Western Hospital Algorithms Brain Stimul. 9 438–52 [PubMed: 26968805]

[3]. Soh D, ten Brinke T R, Lozano AM and Fasano A 2019 Therapeutic window of deep brain stimulation using cathodic monopolar, bipolar, semi-bipolar, and anodic stimulation Neuromodulation Technol. Neural Interface 22 451–5

[4]. Haddock A, Mitchell KT, Miller A, Ostrem JL, Chizeck HJ and Miocinovic S 2018 Automated deep brain stimulation programming for tremor IEEE Trans. Neural Syst. Rehabil. Eng. 26 1618–25 [PubMed: 29994714]

[5]. Sarikhani P, Miocinovic S and Mahmoudi B 2019 Towards automated patient-specific optimization of deep brain stimulation for movement disorders Proc. Annual Int. Conf. IEEE Engineering in Medicine and Biology Society, EMBS (July 2019) pp 6159–62

[6]. Wenzel GR, Roediger J, Brücke C, Marcelino ALDA, Gülke E, Pötter-Nerger M, Scholtes H, Wynants K, Juárez Paz L M and Kühn AA 2021 CLOVER-DBS: algorithm-guided deep brain stimulation-programming based on external sensor feedback evaluated in a prospective, randomized, crossover, double-blind, two-center study J. Parkinsons Dis. 11 1887–99 [PubMed: 34151855]

[7]. Sasaki F, Oyama G, Sekimoto S, Nuermaimaiti M, Iwamuro H, Shimo Y, Umemura A and Hattori N 2021 Closed-loop programming using external responses for deep brain stimulation in Parkinson's disease Parkinsonism Relat. Disord. 84 47–51 [PubMed: 33556765]

[8]. Lorenz R, Monti RP, Violante IR, Anagnostopoulos C, Faisal AA, Montana G and Leech R 2016 The automatic neuroscientist: a framework for optimizing experimental design with closed-loop real-time fMRI Neuroimage 129 320–34 [PubMed: 26804778]

[9]. Stieve BJ, Richner TJ, Krook-Magnuson C, Netoff TI and Krook-Magnuson E 2022 Optimization of closed-loop electrical stimulation enables robust cerebellar-directed seizure control Brain 12 awac051

[10]. Lorenz R, Simmons LE, Monti RP, Arthur JL, Limal S, Laakso I, Leech R and Violante IR 2019 Efficiently searching through large tACS parameter spaces using closed-loop Bayesian optimization Brain Stimul. 12 1484–9 [PubMed: 31289013]

[11]. Boutet A et al. 2021 Predicting optimal deep brain stimulation parameters for Parkinson's disease using functional MRI and machine learning Nat. Commun. 12 3043 [PubMed: 34031407]

[12]. Louie KH, Petrucci MN, Grado LL, Lu C, Tuite PJ, Lamperski AG, MacKinnon CD, Cooper SE and Netoff TI 2021 Semi-automated approaches to optimize deep brain stimulation parameters in Parkinson's disease J. Neuroeng. Rehabil. 18 1–16 [PubMed: 33397401]

[13]. Zhao Z et al. 2021 Optimization of spinal cord stimulation using Bayesian preference learning and its validation IEEE Trans. Neural Syst. Rehabil. Eng. 29 1987–97 [PubMed: 34543198]

[14]. Grado LL, Johnson MD, Netoff TI and Santaniello S 2018 Bayesian adaptive dual control of deep brain stimulation in a computational model of Parkinson's disease PLoS Comput. Biol. 14 e1006606 [PubMed: 30521519]

[15]. Duchet B, Weerasinghe G, Bick C and Bogacz R 2021 Optimizing deep brain stimulation based on isostable amplitude in essential tremor patient models J. Neural Eng. 18 046023 [PubMed: 33821809]

[16]. Duchet B, Weerasinghe G, Cagnan H, Brown P, Bick C and Bogacz R 2020 Phase-dependence of response curves to deep brain stimulation and their relationship: from essential tremor patient data to a Wilson–Cowan model J. Math. Neurosci. 10 1–39 [PubMed: 31993756]

[17]. Basu I, Graupe D, Tuninetti D, Shukla P, Slavin KV, Metman LV and Corcos DM 2013 Pathological tremor prediction using surface EMG and acceleration: potential use in 'ON-OFF' demand driven deep brain stimulator design J. Neural Eng. 10 036019 [PubMed: 23658233]

[18]. Graupe D, Basu I, Tuninetti D, Vannemreddy P and Slavin KV 2013 Adaptively controlling deep brain stimulation in essential tremor patient via surface electromyography Neurol. Res. 32 899–904

[19]. Shukla P et al. 2013 A decision tree classifier for postural and movement conditions in essential tremor patients Int. IEEE/EMBS Conf. Neural Engineering NER pp 117–20

[20]. Khobragade N, Graupe D and Tuninetti D 2015 Towards fully automated closed-loop deep brain stimulation in Parkinson's disease patients: a LAMSTAR-based tremor predictor Proc. Annual Int. Conf. IEEE Engineering in Medicine and Biology Society, EMBS (November 2015) pp 2616–9

[21]. Yamamoto T, Katayama Y, Ushiba J, Yoshino H, Obuchi T, Kobayashi K, Oshima H and Fukaya C 2013 On-demand control system for deep brain stimulation for treatment of intention tremor Neuromodulation Technol. Neural Interface 16 230–5

[22]. Herron JA, Thompson MC, Brown T, Chizeck HJ, Ojemann JG and Ko AL 2016 Chronic electrocorticography for sensing movement intention and closed-loop deep brain stimulation with wearable sensors in an essential tremor patient J. Neurosurg. 127 580–7 [PubMed: 27858575]

[23]. Shahriari B, Swersky K, Wang Z, Adams RP and de Freitas N 2016 Taking the human out of the loop: a review of Bayesian optimization Proc. IEEE 104 148–75

[24]. Fahn S, Tolosa E, Marin C, Jankovic J and Tolosa E 1993 Clinical Rating Scale for Tremor (Baltimore, MD: Williams & Wilkins)

[25]. Herron J and Chizeck HJ 2014 Prototype closed-loop deep brain stimulation systems inspired by Norbert Wiener 2014 IEEE Conf. on Norbert Wiener in the 21st Century: Driving Technology's Future, 21CW 2014—Incorporating the Proc. 2014 North American Fuzzy Information Processing Society Conf., NAFIPS 2014

[26]. Herron J, Denison T and Chizeck HJ 2015 Closed-loop DBS with movement intention Int. IEEE/EMBS Conf. Neural Engineering NER (July 2015) pp 844–7

[27]. Herron JA 2016 Closed-loop deep brain stimulation: bidirectional neuroprosthetics for tremor and BCI

[28]. Rasmussen CE 2004 Gaussian Processes in Machine Learning (Berlin: Springer) pp 63–71

[29]. Matthews AGDG et al. 2017 GPflow: a Gaussian process library using TensorFlow J. Mach. Learn. Res. 18 1–6

[30]. Jones DR, Schonlau M and Welch WJ 1998 Efficient global optimization of expensive black-box functions J. Glob. Optim. 13 455–92

[31]. Knudde N, van der Herten J, Dhaene T and Couckuyt I 2017 GPflowOpt: a Bayesian optimization library using TensorFlow (arXiv:1711.03845)

[32]. Duivenvoorden RRPR, Berkenkamp F, Carion N, Krause A and Schoellig AP 2017 Constrained Bayesian optimization with particle swarms for safe adaptive controller tuning IFAC-PapersOnLine 50 11800–7

[33]. Sui Y, Ch AE, Zurich E, Burdick JW, Krause A and Ch K 2015 Safe exploration for optimization with Gaussian processes Proc. Machine Learning Research (1 June 2015) pp 997–1005

[34]. Wang Z and Jegelka S 2017 Max-value entropy search for efficient Bayesian optimization Proc. Machine Learning Research (17 July 2017) pp 3627–35

[35]. Volkmann J, Moro E and Pahwa R 2006 Basic algorithms for the programming of deep brain stimulation in Parkinson's disease Mov. Disord. 21 S284–9 [PubMed: 16810675]

[36]. Pulliam CL, Heldman DA, Orcutt TH, Mera TO, Giuffrida JP and Vitek JL 2015 Motion sensor strategies for automated optimization of deep brain stimulation in Parkinson's disease Parkinsonism Relat. Disord. 21 378–82 [PubMed: 25703990]

[37]. Heldman DA, Pulliam CL, Urrea Mendoza E, Gartner M, Giuffrida JP, Montgomery EB Jr, Espay AJ and Revilla FJ 2016 Computer-guided deep brain stimulation programming for Parkinson's disease Neuromodulation Technol. Neural Interface 19 127–32

[38]. Noecker AM, Frankemolle-Gilbert AM, Howell B, Petersen MV, Beylergil SB, Shaikh AG and McIntyre CC 2021 StimVision v2: examples and applications in subthalamic deep brain stimulation for Parkinson's disease Neuromodulation 24 248–58 [PubMed: 33389779]

[39]. Jeon H, Lee W, Park H, Lee H, Kim S, Kim H, Jeon B and Park K 2017 Automatic classification of tremor severity in Parkinson's disease using a wearable device Sensors 17 2067

[40]. Rebelo P, Green AL, Aziz TZ, Kent A, Schafer D, Venkatesan L and Cheeran B 2018 Thalamic directional deep brain stimulation for tremor: spend less, get more Brain Stimul. 11 600–6 [PubMed: 29373260]

[41]. Mahlknecht P et al. 2017 Pyramidal tract activation due to subthalamic deep brain stimulation in Parkinson's disease Mov. Disord. 32 1174–82 [PubMed: 28590508]

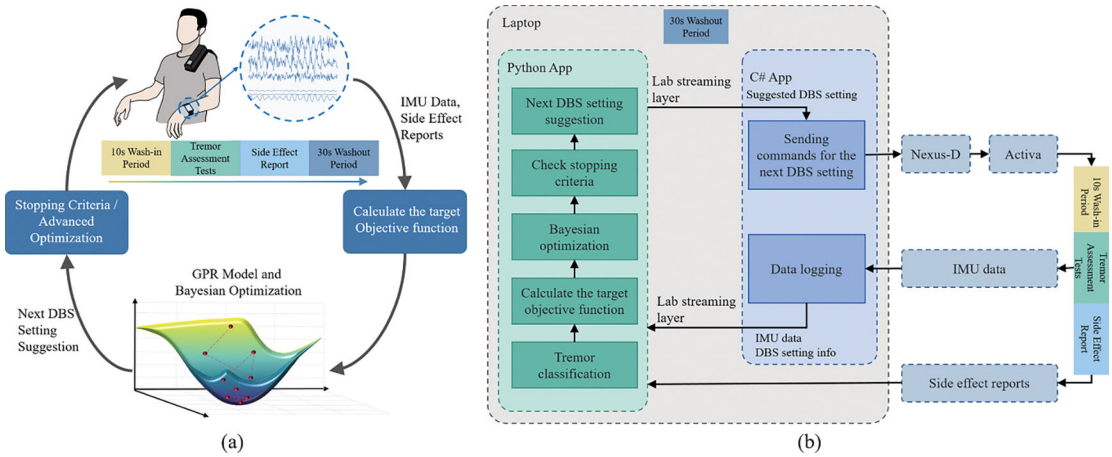

Figure 1.

Overview of the automated DBS optimization framework for tremor programming. (a) After performing the initial baseline tremor evaluation tests without stimulation, at each iteration, the software automatically sets the next DBS setting to be tested followed by 10 s wash-in period, followed by tremor evaluation tests each for 10 s. The recorded inertial measurement unit (IMU) data and side-effect reports are used to update the surrogate gaussian process regression (GPR) model and optimizer suggests the next best sample to be tested. Before evaluating the next suggested DBS setting, the stopping criteria module determines whether the optimum has been found or advanced stimulation is needed. (b) A detailed schematic demonstrating the software design of the automated DBS programming system. The software application receives IMU data over a Bluetooth connection from the smartwatch, as well as side effects reported by the patient through a graphical user interface and send the information to the Python section of the application. The calculation of the objective measure (surrogate function) and choice of the next DBS setting (acquisition function) are handled within the Python section. The C# software application receives the stimulations settings from the Python application and sends stimulation commands to the Nexus-D, which communicates with patient's implanted Activa IPG.

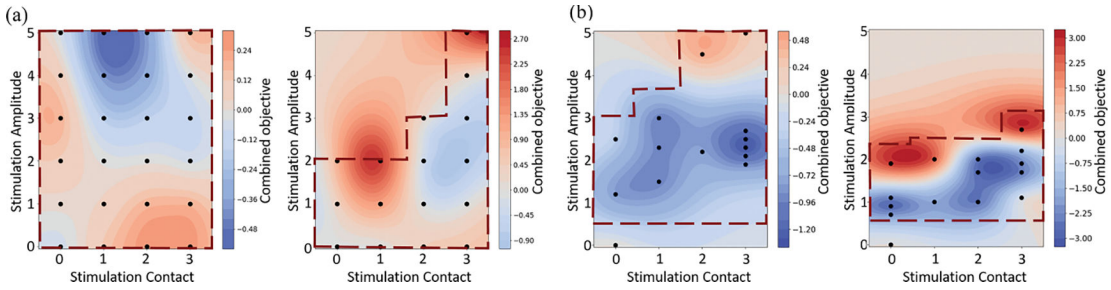

Figure 2.

GPR model mean surface of the combined objective measure (including baseline-subtracted watch tremor score and side effect score) varies across patients. (a) Mean surfaces for two patients with grid-search sampling strategy from a prior study [4]. The sampling resolution is 1 V amplitude increments. (b) Mean surfaces for two patients from the current study with sampling using Bayesian optimization that evaluates more samples in areas with greater chance of tremor improvement and with a finer resolution (0.2 V amplitude increments). The surfaces are color-coded with the value of the combined objective measure where blue shows negative objective values reflecting tremor improvement compared to baseline either without or with mild side effect and red shows positive values reflecting that DBS settings are not effective or side effects are pronounced. The black circles represent sampled DBS settings during the automated DBS optimization. The red dashed lines show the clinician-defined safe exploration boundaries of the parameter space.

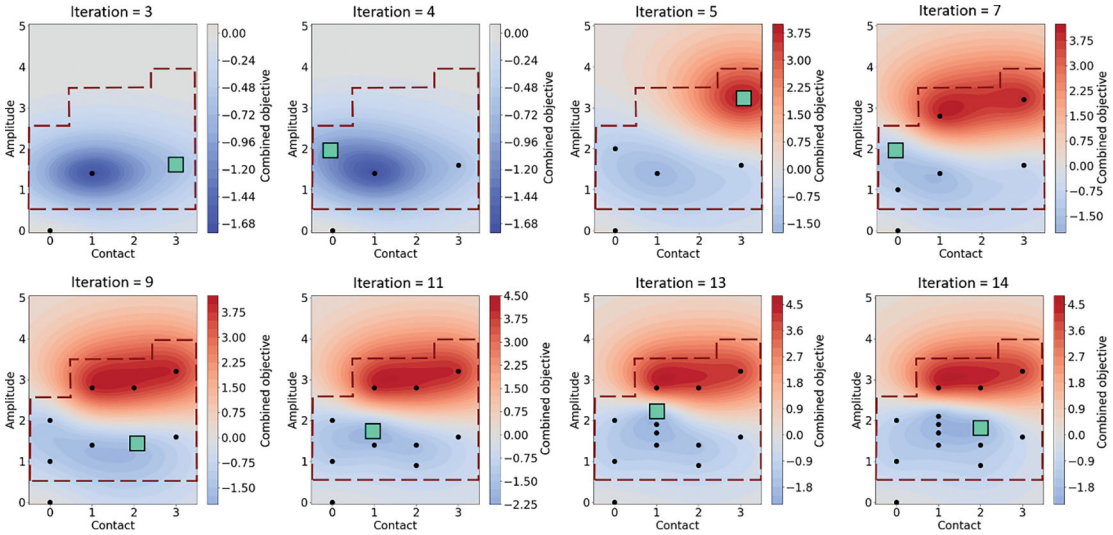

Figure 3.

Example of patient-specific adaptive sampling of Bayesian optimization (patient 02). Each panel shows the mean surface of the GPR model that updates after each iteration. The value of the combined objective measure is color-coded. The dashed dark red lines demonstrate the clinician-defined maximum tolerable exploration boundaries. The black circles show the previously collected samples and the green square show the sample being tested at each iteration. The black circle outside the red dashed lines at (0, 0) demonstrates the baseline, where the patient's IPG was inactive. The sample suggestions are automated by the DBS optimization framework. Samples are more densely distributed around the more promising regions of the parameter space (more tremor improvement with fewer side effects). This adaptive behavior of the DBS optimization framework makes it patient-specific; that is the samples are adaptively suggested based on the patient's response at previous iterations.

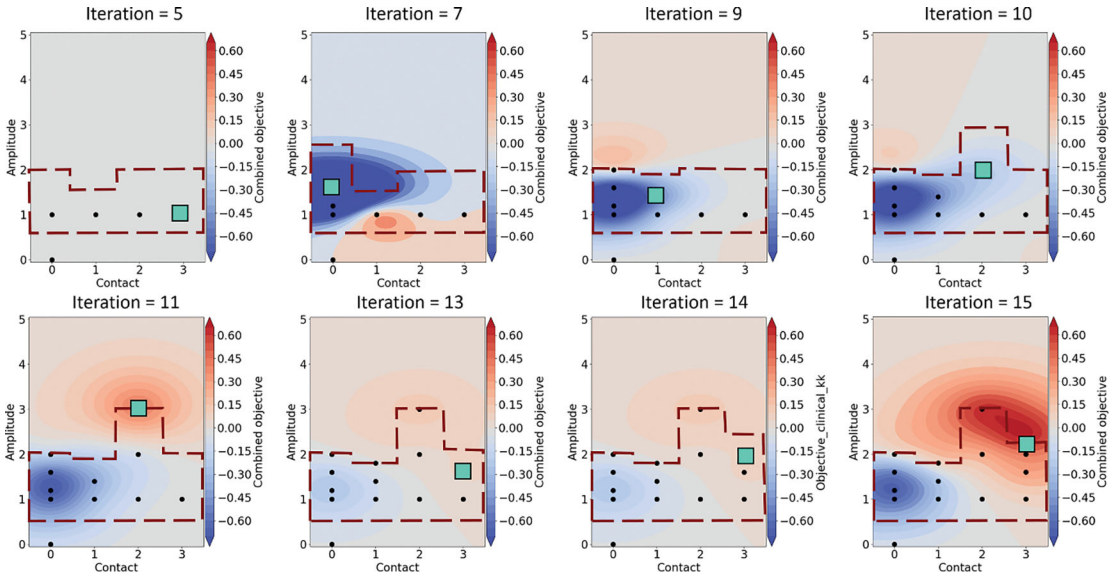

Figure 4.

Safe Bayesian optimization during phase II of the experiments (patient 14). Selected iterations during monopolar stimulation. The mean surface of the GPR model and safe stimulation exploration boundaries (dashed lines) update as more data are collected at each iteration. The value of the combined objective measure is color-coded. The black circles represent collected samples and the green square is the current sample being tested at each iteration. The black circles outside the red dashed lines at (0, 0) demonstrates the baseline, where the patient's IPG was inactive.

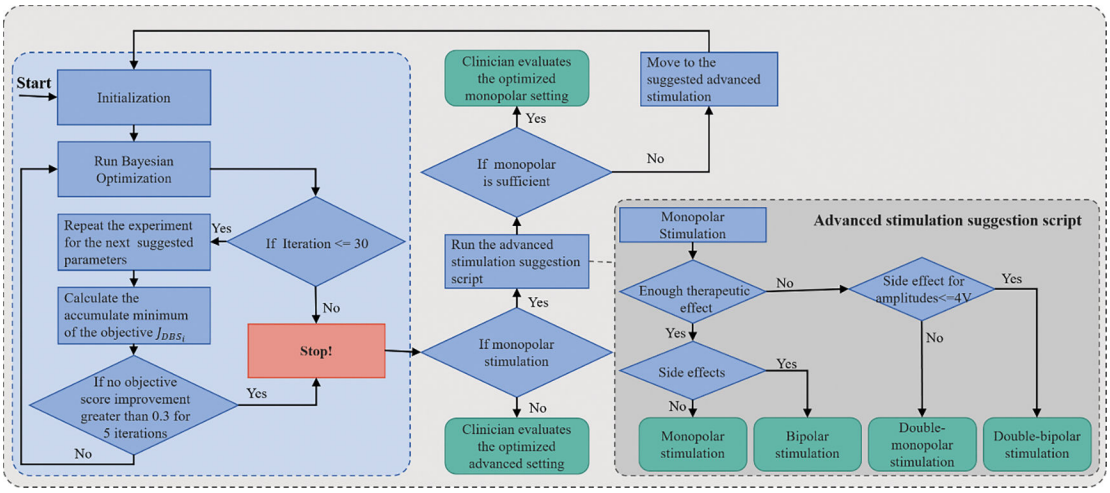

Figure 5.

High-level schematic of the decision-making process of the automated DBS optimization framework. The darker gray area is the schematic of the advanced optimization suggestion algorithm modeled after the clinical decision-making process.

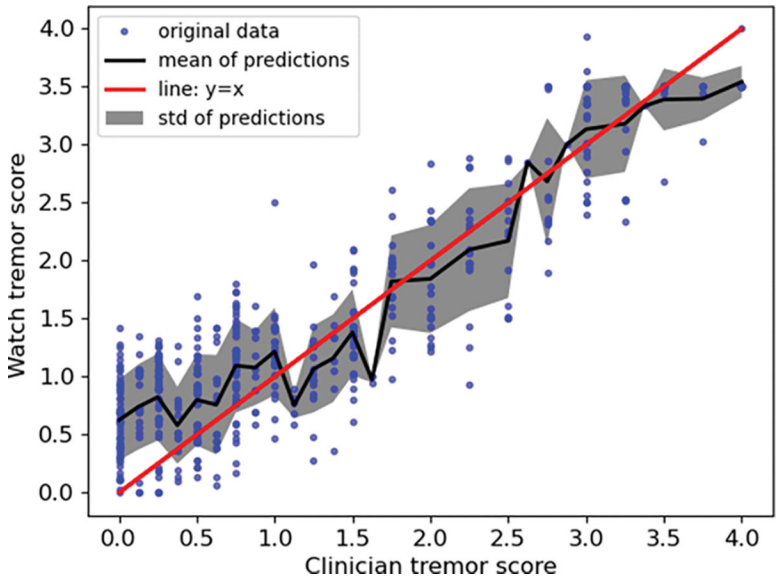

Figure 6.

Additional validation of tremor score classifier [4]. Blue dots represent the average watch tremor scores plotted against the average clinician tremor score for selected tremor assessment tasks (rest, arms extended, arms flexed, finger-to-nose motion). Each dot represents one DBS setting that was tested during the experiments. The black solid line and the gray shaded area show the mean and standard deviation of the watch tremor scores. The red solid line is the line $y = x$ and the r-squared value of the fit to the $y = x$ line is 0.69.

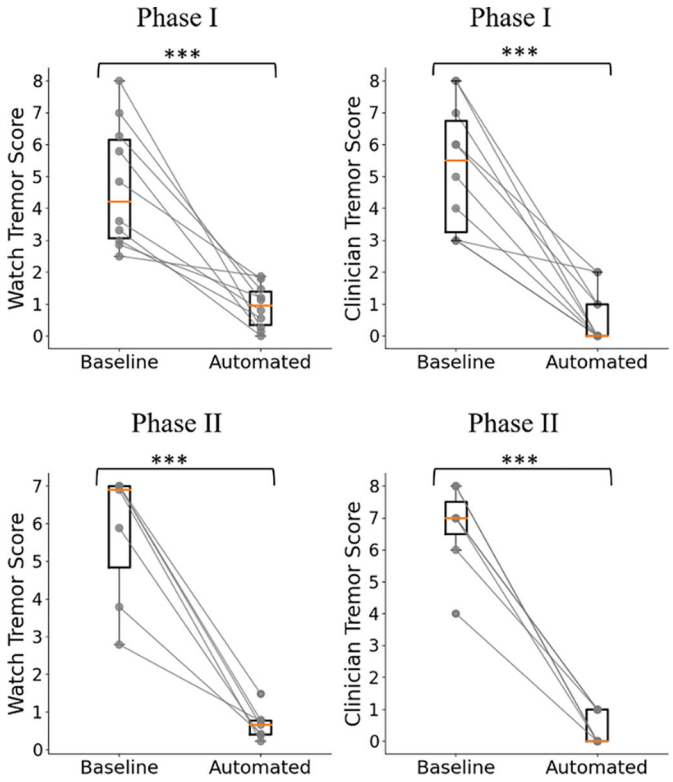

Figure 7.

Clinical efficacy of automated DBS programming. Comparison of the patients' tremor severity scores at baseline stimulation off condition and the optimal automated setting measured by the watch (left column) and the optimal automated setting scored by a blinded clinician (right column). Top row refers to phase I (clinician-defined safe amplitudes), and bottom row to phase II (safe Bayesian optimization algorithm) experiments. The asterisk shows the conditions with statistically significant difference ($^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$). The tremor score is the sum of two tremor assessment tasks utilized during the automated DBS optimization session (max 8).

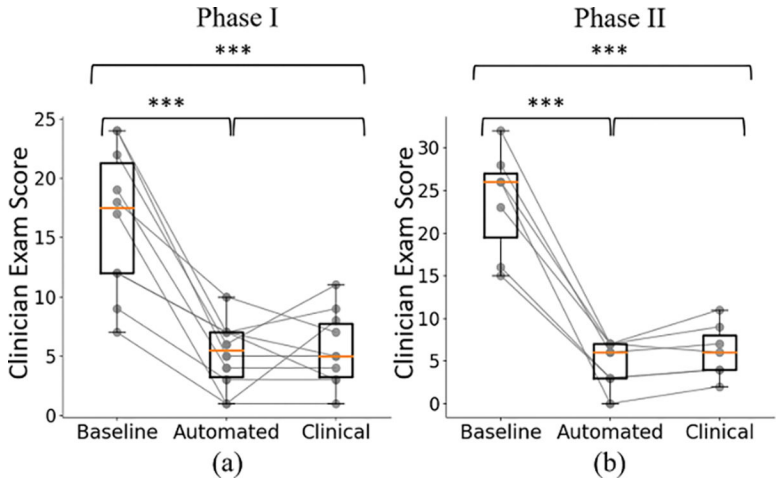

Figure 8.

Clinical efficacy of automated DBS programming compared to clinical setting. Comparison of the patients' tremor severity scores at baseline (no stimulation), the best automated setting, and previously established best clinical setting during phase I (a) and phase II (b) based on the clinician scores during the comprehensive clinical exam. The comprehensive exam included the following items from FTM tremor scale: rest, arms extended, arms flexed, and finger-to-nose motion arm tremor contralateral to DBS lead, handwriting (if dominant hand tested), two spiral drawings, and line drawing. Both the patient and clinician were aware of the stimulation condition. The asterisk shows the conditions with statistically significant difference ($^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$).

Table 1.

Clinical characteristics and automated programming experiment outcome during the phase I of the experiments using the clinician-defined maximum safe boundaries of the parameter space.

Patient ID	Age, sex	Dx	DBS target	Time since DBS surgery (months)	Time since selection of clinical setting (months)	Tremor score[a] at baseline	Clinical DBS setting	Tremor score[a] at clinical setting	Automated DBS setting	Tremor score[a] at automated setting	Number of tested DBS settings	Patient preference
1	63, M	PD	R STN	34	28	7	C + 3−, 3.5 V, 90 μs, 130 Hz	1	C + 3−, 2.3 V, 90 μs, 130 Hz	1	Monopolar = 15, Advanced = NA	Same
2	71, F	PD	L STN	42	12	17	2–3+, 4.6 V, 90 μs, 160 Hz	8	C + 1−, 1.9 V, 90 μs, 160 Hz	1	Monopolar = 17, Advanced = NA	Same
3	61, M	PD	R STN	63	17	12	C + 2−, 3.5 V, 60 μs, 165 Hz	9	C + 1−, 0.7 V, 60 μs, 165 Hz	7	Monopolar = 15, Advanced = NA	Automated
4	82, F	ET	L VIM	54	1	24	1–3+, 2.4 V, 90 μs, 130 Hz	3	C + 1−, 1.1 V, 90 μs, 130 Hz	7	Monopolar = 15, Advanced = NA	Clinical
5	79, M	ET	R VIM	54	1	22	9–11+, 2.5 V, 60 μs, 190 Hz	11	C + 10−, 2.3 V, 60 μs, 190 Hz	6	Monopolar = 14, Advanced = NA	Automated
6	76, M	ET	L VIM	37	11	9	1–3+, 2.5 V, 90 μs, 170 Hz	3	1–2+, 3.5 V, 90 μs, 170 Hz	3	Monopolar = 15, Advanced = 14	Same
7	77, M	ET	L VIM	43	1	18	1–2–0+, 4.6 V, 60 μs, 170 Hz	7	C + 1−, 1.1 V, 60 μs, 170 Hz	10	Monopolar = 15, Advanced = NA	Clinical
8	76, F	ET	L VIM	63	1	19	2–3+, 3.6 V, 90 μs, 160 Hz	4	2–1+, 4 V, 90 μs, 160 Hz	4	Monopolar = 15, Advanced = 15	Same
9	69, M	PD	R STN	70	5	12	0–1–2+, 4 V, 60 μs, 140 Hz	5	1–0+, 3.7 V, 60 μs, 140 Hz	7	Monopolar = 15, Advanced = 15	Clinical
10	57, M	PD	R STN	42	1	24	0–3+, 4.7 V, 90 μs, 130 Hz	5	2–1+, 4.7 V, 90 μs, 130 Hz	5	Monopolar = 15, Advanced = 15	Same

[a]Clinician administered FTM tremor scale subset including rest, postural (extended and flexed), and action arm tremor contralateral to optimized DBS lead, handwriting (if dominant hand tested), and drawings (max score 28–32).

Table 2.

Clinical characteristics and automated programming experiment outcome phase II with automated discovery of the safe parameter space using safe Bayesian optimization algorithm.

Patient ID	Age, sex	Dx	DBS target	Time since DBS surgery (months)	Time since selection of clinical setting (months)	Tremor score[a] at baseline	Clinical DBS setting	Tremor score[a] at clinical setting	Automated DBS setting	Tremor score[a] at automated setting	Number of tested DBS settings	Patient preference
11	67, F	ET	R VIM	108	2	16	0 + 1–3+, 5.2 V, 90 μs, 150 Hz	4	3–2+, 4.8 V, 90 μs, 150 Hz	3	Monopolar = 13, Advanced = 13	Automated
10	57, M	PD	R STN	45	3	28	0–3+, 4.7 V, 90 μs, 130 Hz	7	2–3+, 4.8 V, 90 μs, 130 Hz	6	Monopolar = 17, Advanced = 13	Same
12	60, M	PD	L STN	49	21	23	0–1–3+, 4.9 V, 90 μs, 180 Hz	6	C + 0–1−, 2.2 V, 60 μs, 180 Hz	7	Monopolar = 28, Advanced = 19	Same
5	79, M	ET	L VIM	58	4	26	2–3+, 3.7 V, 120 μs, 190 Hz	9	2–1+, 3.8 V, 120 μs, 190 Hz	7	Monopolar = 13, Advanced = 13	Clinical
13	61, M	PD	R STN	49	6	26	1–2+, 4.4 mA, 90 μs, 130 Hz	2	3–1+, 4 V, 90 μs, 130 Hz	0	Monopolar = 21, Advanced = 15	Automated
14	85, M	PD	L VIM	33	2	32	0–2+, 2.8 V, 90 μs, 130 Hz	11	0–2+, 3.6 V, 90 μs, 130 Hz	7	Monopolar = 15, Advanced = 14	Automated
15	68, M	PD	L STN	63	3	15	C + 3−, 2.6 V, 60 μs, 150 Hz	4	C + 2−, 1.6 V, 60 μs, 150 Hz	3	Monopolar = 17, Advanced = NA	Same

[a]Clinician administered FTM tremor scale subset including rest, postural (extended and flexed), and action arm tremor contralateral to optimized DBS lead, handwriting (if dominant hand tested), and drawings (max score 28–32).