2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Safe Explorative Bayesian Optimization – Towards Personalized Treatments in Plasma Medicine

Kimberly J. Chan, Joel A. Paulson, and Ali Mesbah

***Abstract*— This paper considers the problem of Bayesian optimization (BO) for systems with safety-critical constraints. Recent work has shown that a theoretically consistent way to account for constraints in BO is to relax the constraint functions such that the feasible region has a high probability of containing the global solution. However, by construction, these approaches are unable to ensure safe/feasible operation at every query, which is unacceptable in safety-critical applications. Alternatively, safe BO methods force the query points to remain in the interior of a partially-revealed safety region, which may result in unacceptable (and unquantified) performance losses. This paper presents a new safe BO method that avoids these performance losses by systematically incorporating potential performance gains from enlargement of the safety region. The proposed method avoids getting stuck at suboptimal points based on a potentially small initial safety region due to limited initial exploration of the safety boundary. The performance of the proposed method is demonstrated for safe control of a cold atmospheric plasma jet towards personalized plasma medicine.**

I. Introduction

Advances in data-driven control and decision-making capabilities have created significant opportunities for autonomy for vehicles, robots, and biomedical devices [1]. Interactions with humans make safety a fundamental requirement for these autonomous systems. Thus, the underlying problem for autonomy can be generally posed as a constrained optimization problem of the form

$$\min_{x\in\mathcal{X}} \left\{ f^0(x) : f^i(x) \geq 0,\ \forall i = 1,\dots,m \right\}, \tag{1}$$

where x are the decision variables (i.e., modifiable parameters), $f^0 : \mathcal{X} \to \mathbb{R}$ is the objective function, $f^i : \mathcal{X} \to \mathbb{R}$ are constraints, and $\mathcal{X} \subset \mathbb{R}^{n_x}$ is some compact domain.

In autonomous systems, it is often the case that the mathematical structure of the objective $f^0(x)$ and constraints $f^i(x)$ are not exactly known, e.g., those derived from closed-loop trajectories. In such cases, we often refer to the functions $\{f^i(x)\}_{i=0}^m$ as "black-box" in the sense that they can only be learned from noisy observations at specific query points $x \in \mathcal{X}$. These observations must then be used in a strategy to compute an optimum for x. One such strategy is Bayesian optimization (BO) [2]. BO is a sequential decision-making strategy that uses probabilistic surrogate models of $\{f^i(x)\}_{i=0}^m$ to optimize a proxy problem to (1). The surrogate models, often represented with Gaussian Processes (GPs), are updated via Bayesian inference [3]. Optimizing the proxy problem is facilitated by an acquisition function $\alpha(x) : \mathcal{X} \to \mathbb{R}$, which, in some form, leverages the uncertainty in the posterior model. This feature of optimizing with respect to both the belief of the optimum and the uncertainty surrounding this belief is commonly referred to as the trade-off between exploration and exploitation. By iteratively querying and updating the surrogates, BO systematically explores the design space $\mathcal{X}$ to find an optimum.

For the safe interaction of autonomous systems with humans, it is imperative to ensure that the proposed parameter choices satisfy constraints $\{f^i(x)\}_{i=1}^m$. In other words, evaluating any arbitrary x in $\mathcal{X}$ can lead to constraint-violating designs that can have dangerous outcomes. Yet, identifying a safe, or feasible, design space $\mathcal{F} \subseteq \mathcal{X}$ when the constraints are unknown is challenging. To this end, several recent works have been proposed in the realm of safe BO. At its core, the safe BO problem is exactly the constrained optimization problem (1) with strict adherence to the constraints. Two common approaches are taken: (i) a penalty-based strategy, where the constraints act as a penalty term in the objective; and (ii) a safe set strategy, where points are only chosen from an estimation of the safe region. In [4], the acquisition function is augmented with barrier functions, a take on (i); it uses the posterior estimates of the constraints to directly penalize the acquisition objective to limit the search to revealed safe points. Here, the solution will only be locally optimal near the initial safe design point x_0. Alternatively, [5] uses the posterior models to compute a partially-revealed safe set using Lipschitz continuity properties. This safe set is further subdivided into a set of potential "optimizers" and a set of potential "expanders." Then, the most uncertain element from the combination of the optimizers and expanders is suggested as the next query. This method takes an exploration-first perspective to maximize the discovery of safe points and then switches focus to the standard exploitation/exploration trade-off. This method may lead to "wasteful" queries within the feasible region.

In an attempt to expand the feasible region, [6] proposes an ϵ-greedy approach to switch to a pure exploratory strategy to directly explore the boundary of the feasible region. However, this exploratory procedure remains agnostic to improvements in the main objective. This paper presents SEBO (Safe Explorative Bayesian Optimization), a new safe BO method that avoids potential performance losses by incorporating information gained by expanding the feasible

K. J. Chan and A. Mesbah are with the Department of Chemical and Biomolecular Engineering, University of California, Berkeley, CA 94720, USA. {kchan45,mesbah}@berkeley.edu

J. A. Paulson is with the Department of Chemical and Biomolecular Engineering, The Ohio State University, Columbus, OH 43210, USA. paulson.82@osu.edu

This work was supported by the US NSF under Grants 2130734 and 2237616.

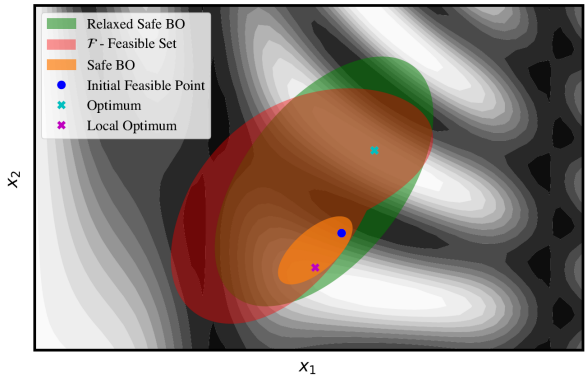

Fig. 1. Exemplary feasible set using safe Bayesian optimization (BO) (in orange) versus a relaxed problem (in green). The true feasible set is in red, while the initial feasible point is in blue. The true optimum is denoted with a cyan "x", and a local optimum is denoted with a magenta "x". Contours of the objective are in gray-scale with more optimal spaces in white.

region. As depicted in Fig. 1, penalty-based strategies for safe BO may be prone to being overly conservative such that they may get stuck in the locally feasible region near the initial safe point. SEBO uses a relaxed formulation to widen the search space that more likely encapsulates the true optimum. Safety is ensured by projecting back to the estimated safe region, but at the same time, maximizing the potential to increase knowledge around the safe set in the direction of improvement. Thus, SEBO effectively incorporates directed information to explore the safe region(s).

We demonstrate SEBO for an example application for personalized plasma medicine, which is an emerging field involving the use of cold atmospheric plasmas (CAPs) for a variety of medical treatments [7]. Automated CAP treatments using advanced control (e.g., model predictive control, MPC) are necessary for ensuring effective delivery of plasma effects [8], [9]. Tailoring the plasma effects is key to ensuring the efficacy of plasma treatments [10]. However, the underlying mechanisms of plasma-surface interactions can only be quantified for a population of subjects [8], [11]. Therefore, iterative improvements in delivery of plasma effects using BO will enable personalization of CAP treatments, wherein ensuring patient safety is of the utmost importance. We compare SEBO's performance to alternative safe BO approaches and demonstrate that it can mitigate getting stuck in a local feasible region while realizing safe CAP treatments.

II. Safe Bayesian Optimization using Logarithmic Barrier Functions

Since the model and constraint functions in (1) are unknown, we must learn them from data. Here, we focus on the case that these functions can be modeled as independent Gaussian processes (GPs)

$$f^i(x) \sim \mathcal{GP}(\mu^i(x), k^i(x,x')),\ \forall i = 0,\dots,m, \tag{2}$$

where $\mu^i(x) = \mathbb{E}\{f^i(x)\}$ denotes the prior mean function and $k^i(x,x') = \mathbb{E}\{(f^i(x)-\mu^i(x))(f^i(x')-\mu^i(x'))\}$ denotes the prior covariance (kernel) function of the objective and safety constraints. GP models are non-parametric and have the property that the posterior model, conditioned on noisy observations $\boldsymbol{y}_n^i = (y_1^i,\dots,y_n^i)$ at query points $(x_1,\dots,x_n)$, remains a GP with the following analytic expressions for the mean, kernel, and standard deviation functions

$$\mu_n^i(x) = \boldsymbol{k}_n^{i\top}(x)(\boldsymbol{K}_n^i + \eta^i I)^{-1}\boldsymbol{y}_n^i, \tag{3a}$$
$$k_n^i(x,x') = k^i(x,x') - \boldsymbol{k}_n^{i\top}(x)(\boldsymbol{K}_n^i + \eta^i I)^{-1}\boldsymbol{k}_n^i(x'),$$
$$\sigma_n^i(x) = \sqrt{k_n^i(x,x)}, \tag{3b}$$

where $\boldsymbol{k}_n^i(x) = [k^i(x_1,x),\dots,k^i(x_n,x)]^\top$, $\boldsymbol{K}_n^i$ is the positive definite kernel matrix whose elements are given by $[\boldsymbol{K}_n^i]_{\nu,\omega} = k_i(x_\nu,x_\omega)$ for all $\nu,\omega \in \{1,\dots,n\}$, and $\eta^i > 0$ is the variance of a zero-mean Gaussian noise model for the observations, i.e., $y_j^i = f^i(x_j) + \epsilon_j^i$ for some ϵ_j^i that is R-sub Gaussian noise [12].

If the GP models are sufficiently "well-calibrated," then they can provide high probability confidence bounds on $\{f^i(x)\}_{i=0}^m$. We summarize this requirement below.

Assumption 1 (Well-calibrated GPs): The GP models for the unknown objective and constraint functions $\{f^i(x)\}_{i=0}^m$ satisfy the inequality below $\forall x \in \mathcal{X}$, $n \geq 0$, and $i = 0,\dots,m$

$$|f^i(x) - \mu_n^i(x)| \leq \sqrt{\beta_{n+1}^i}\sigma_n^i(x), \tag{4}$$

with probability at least $1-\delta$ for some $\delta \in (0,1)$.

This assumption can be satisfied by properly selecting the sequence of confidence bound parameters $\{\beta_{n+1}^i\}_{n\geq 0}$ as long as the functions $\{f^i(x)\}_{i=0}^m$ have a bounded reproducing kernel Hilbert space (RKHS) norm; see [12] for expressions for $\{\beta_{n+1}^i\}_{n\geq 0}$, which are defined in terms of the maximum information gain.

For convenience, we rewrite Assumption 1 in terms of the lower confidence bound (LCB) and upper confidence bound (UCB) on the unknown functions at iteration n

$$l_n^i(x) = \mu_n^i(x) - \sqrt{\beta_{n+1}^i}\sigma_n^i(x), \tag{5a}$$
$$u_n^i(x) = \mu_n^i(x) + \sqrt{\beta_{n+1}^i}\sigma_n^i(x), \tag{5b}$$

where the inequality (4) can be equivalently stated as $f^i(x) \in [l_n^i(x), u_n^i(x)]$ using the LCB/UCB definitions. The main idea behind a safe BO procedure is then to sequentially select new sample points $x_1, x_2, \dots$ that have a high probability of satisfying safety constraints at every iteration. This differs from standard BO that would query by solving $x_{n+1} \in \arg\min_{x\in\mathcal{X}} l_n^0(x)$ (using an LCB acquisition function). As proposed in [4], one can ensure x_{n+1} remains in the interior of the partially-revealed safe set by solving

$$x_{n+1} \in \arg\min_{x\in\mathcal{X}} \left\{ l_n^0(x) + \tau \textstyle\sum_{i=1}^m \mathcal{B}_{l_n^i}(x) \right\}, \tag{6}$$

where $\mathcal{B}_g(x) = -\ln(g(x))$ is the logarithmic barrier applied to a constraint $g(x) \geq 0$ and $\tau > 0$ is a tunable parameter that ensures the barrier term converges to the exact indicator penalty function in the limit $\tau \to 0$. Notice that (6) accounts for both performance and safety. The safety guarantees conferred by (6) are summarized in the following theorem.

Theorem 1 (Safe Learning [4]): Let Assumption 1 hold, the feasible set $\mathcal{F} = \{x \in \mathcal{X} : f^i(x) \geq 0, \forall i = 1, \ldots, m\}$ be non-empty, and there exists at least one known safe starting point $x_0 \in \mathcal{F}$. Then, the sequence of query points $\{x_n\}_{n\geq 1}$ generated by (6) satisfies

$$\Pr\left\{f^i(x_n) \geq 0,\ \forall i = 1, \ldots, m, \forall n \geq 1\right\} \geq 1 - \delta, \quad (7)$$

for any chosen $\delta \in (0, 1)$.

The proof of this theorem is based on three key arguments. First, the partially-revealed feasible region, defined by

$$\hat{\mathcal{F}}_n = \{x \in \mathcal{X} : l_n^i(x) \geq 0,\ \forall i = 1, \ldots, m\}, \quad (8)$$

must be contained within the true feasible region $\hat{\mathcal{F}}_n \subseteq \mathcal{F}$ with high probability. Second, $\hat{\mathcal{F}}_n \neq \emptyset$ must be non-empty given a known safe point x_0. Third, the log-barrier term in (6) always guarantees the next query point x_{n+1} is contained in this estimated region, i.e., $x_{n+1} \in \hat{\mathcal{F}}_n$.

The primary challenge encountered by this type of safe BO approach is that it does not have any direct incentive to grow the size of the partially-revealed safety region $\hat{\mathcal{F}}_n$. Thus, it may become "stuck" in the sense that $\hat{\mathcal{F}}_n \subset \mathcal{F}$ as $n \to \infty$, which could lead to sub-optimal performance in cases where the global solution to (1) satisfies $x^\star \in \mathcal{F} \setminus \hat{\mathcal{F}}_n$. We look to overcome this challenge by introducing SEBO.

III. Safe Explorative Bayesian Optimization

The main idea motivating the proposed SEBO method is that there is value in enlarging the certifiable safety region at every iteration of BO to ensure that we do not get stuck in a sub-optimal solution. Thus, we require a metric that, when evaluated at any $x \in \mathcal{X}$, provides a reasonable measure of the potential benefit of querying the constraints at that point in the future. Let $\text{Vol}(\hat{\mathcal{F}}_n)$ denote the volume of the partially-revealed safe set. We can propose the following safety-based acquisition function that we refer to as the expected safety improvement (ESI)

$$\text{ESI}_n(x) = \mathbb{E}_n\left\{\text{Vol}(\hat{\mathcal{F}}_{n+1}) - \text{Vol}(\hat{\mathcal{F}}_n) \mid x_{n+1} = x\right\}, \quad (9)$$

where $\mathbb{E}_n\{\cdot\}$ denotes the expectation with respect to the posterior distribution given all function evaluations up until iteration n. There are two important challenges with (9): (i) it is expensive to compute and optimize since it requires repeated estimation of the volume of a set, though this can in principle be done with Monte Carlo methods (see, e.g., [13]); and (ii) any growth in the safety region is valued by this metric, which can impede the ability to discover new safe points that are likely to improve performance over multiple steps in the future. SEBO attempts to overcome both of these challenges by applying a series of steps that do not compromise the safety guarantees established in Theorem 1.

The first major step of SEBO is to decide when the choice in (6) is likely to contain low information content. The most straightforward approach is to check if $\sigma_n^0(x_{n+1}) \leq \varepsilon$, where $\varepsilon \geq 0$ is a user-specified tolerance value. This implies we can confidently predict the value of $f^0(x_{n+1})$ and, thus, have no additional room for improvement within $\hat{\mathcal{F}}_n$. Whenever such a situation occurs, we must explore outside the current safe region. To decide a new search direction, we solve the following relaxed problem

$$x_{n+1}^r \in \arg\min_{x\in\mathcal{X}} \left\{l_n^0(x) + \tau \textstyle\sum_{i=1}^m \mathcal{B}_{u_n^i}(x)\right\}, \quad (10)$$

where the LCB in the log-barrier term in (6) is replaced with the UCB. This change fundamentally alters the way that constraints are handled in the search process. In particular, (10) operates over a relaxed feasible region

$$\tilde{\mathcal{F}}_n = \{x \in \mathcal{X} : u_n^i(x) \geq 0,\ \forall i = 1, \ldots, m\}. \quad (11)$$

Since $\tilde{\mathcal{F}}_n$ contains the global solution with high probability under Assumption 1 [14], sampling at $\{x_n^r\}_{n\geq 1}$ will guarantee convergence to the global solution; however, it will result in loss of the safety guarantees. Instead of directly sampling at this point, SEBO finds the closest safe point to x_{n+1}^r by solving the following projection problem

$$x_{n+1}^p \in \arg\min_{x\in\hat{\mathcal{F}}_n} \|x - x_{n+1}^r\|. \quad (12)$$

For any $x_{n+1}^r \notin \hat{\mathcal{F}}_n$, the projected point $x_{n+1}^p \in \partial\hat{\mathcal{F}}_n$ will lie on the boundary of the partially-revealed safe region and, thus, is more likely to expand the boundary. However, this projection does not account for the uncertainty of the constraint functions and may sample low uncertainty points. Thus, the final step of SEBO is to find the point with the largest sum of standard deviations in a neighborhood of the safe region boundary around x_{n+1}^p, i.e.,

$$x_{n+1}^s \in \arg\max_{x\in\partial\hat{\mathcal{F}}_n \cap N_b(x_{n+1}^p)} \textstyle\sum_{i=1}^m \sigma_n^i(x), \quad (13)$$

where $N_b(z) = \{x : \|x - z\| \leq b\}$ is a b-radius ball around point z. When the solution to (6) has low information, x_{n+1}^s is proposed as the new query point. The SEBO method is summarized in Algorithm 1. Although the practical performance of SEBO will be affected by the choice of parameters ε and b, they will not affect the safety properties, as summarized below.

Algorithm 1 Safe Explorative Bayesian Optimization

Input(s): Domain $\mathcal{X}$, safe point $x_0 \in \mathcal{F}$, initial data $\mathcal{D}_0 = \{x_0, \{f^i(x_0)\}_{i=0}^m\}$, $m+1$ GP models (2), confidence bound parameters $\{\beta_{n+1}^i\}_{n\geq 0}$, barrier parameter $\tau > 0$, switching tolerance $\varepsilon \geq 0$, and exploration radius $b \geq 0$.

for $n = 0, 1, \ldots$ **do**
  $x_{n+1} \leftarrow \arg\min_{x\in\mathcal{X}}\{l_n^0(x) + \tau\sum_{i=1}^m \mathcal{B}_{l_n^i}(x)\}$
  if $\sigma_n^0(x_{n+1}) \leq \varepsilon$ **then**
    $x_{n+1}^r \leftarrow \arg\min_{x\in\mathcal{X}}\{l_n^0(x) + \tau\sum_{i=1}^m \mathcal{B}_{u_n^i}(x)\}$
    $x_{n+1}^p \leftarrow \arg\min_{x\in\hat{\mathcal{F}}_n} \|x - x_{n+1}^r\|$
    $x_{n+1} \leftarrow \arg\max_{x\in\partial\hat{\mathcal{F}}_n\cap N_b(x_{n+1}^p)} \sum_{i=1}^m \sigma_n^i(x)$
  end if
  Query x_{n+1} and observe objective and constraints
  Update data $\mathcal{D}_{n+1} \leftarrow \mathcal{D}_n \cup \{x_{n+1}, \{f^i(x_{n+1})\}_{i=0}^m\}$
  Update GP models with $\mathcal{D}_{n+1}$ using (3)
end for

Theorem 2: Let the assumptions in Theorem 1 hold. Then, for any choice of $\varepsilon, b \geq 0$, the sequence of query points $\{x_n^s\}_{n\geq 1}$ generated by (13) will satisfy the safety constraints (7), with x_n replaced by x_n^s, for any $\delta \in (0,1)$.

Proof: The projection (12) and exploration (13) steps of SEBO ensure $x_{n+1}^s \in \hat{\mathcal{F}}_n$ for all $n \geq 0$ such that the same arguments used in the proof of Theorem 1 follow. ■

It is interesting to note that Algorithm 1 reduces to the safe BO method in [4] in the case that $\sigma_n^0(x_{n+1}) > \varepsilon$ always holds, which is guaranteed to be true whenever $\varepsilon = 0$. As such, we can interpret SEBO as a generalization of this method. SEBO will be particularly useful whenever one starts with a very restrictive inner approximation of $\mathcal{F}$.

IV. Personalized Plasma Treatment Guidance

A. CAP Jet Modeling and Control

We consider a kHz-excited CAP jet (CAPJ) in helium [9]. The manipulated inputs are applied power P (in Watts) and helium flow rate q (in standard liters per minute, SLM). The measured outputs are maximum surface temperature T (°C) and total optical intensity I (in arbitrary units) of the plasma at the plasma-surface incident point. The system dynamics $h(\cdot)$ are described by an observable, canonical form of a linear time-invariant model

$$s(k+1) = As(k) + Ba(k) + w(k), \tag{14}$$

where k is the discrete-time step, $s = [T, I]^\top \in \mathbb{R}^2$ is the vector of states, $a = [P, q]^\top \in \mathbb{R}^2$ is the vector of manipulated inputs, w is a stochastic variable that represents the overall system uncertainty, and A, B are the state-space matrices identified using subspace identification [15].

CAP treatments rely on the quantification of the delivered plasma effects to a surface. We describe the accumulation of thermal effects on a target with a metric called cumulative equivalent minutes (CEM) [16], [17] given by

$$\mathrm{CEM}(k+1) = \mathrm{CEM}(k) + K^{(T_{\mathrm{ref}} - T(k))}\delta t, \tag{15}$$

where $K > 0$ is an exponential base dependent on physical properties of the substrate, $T_{\mathrm{ref}} = 43$°C is the reference temperature, δt is the sampling time, and $\mathrm{CEM}(0) = 0$. Here, $\delta t = 0.5$ s in accordance with our open-loop data collection.

For a controlled plasma treatment, we use MPC, which is formulated in terms of the optimal control problem

$$\begin{aligned}
\min_{\mathbf{s}(k),\mathbf{a}(k)} \quad & (\mathrm{CEM}_{sp} - \mathrm{CEM}_c(N_p|k))^2 && \text{(16a)}\\
\text{s.t.} \quad & s(i+1|k) = h_c\left(s(i|k), a(i|k)\right), && \text{(16b)}\\
& (s(i|k), a(i|k)) \in \mathcal{S} \times \mathcal{A}, && \text{(16c)}\\
& s(0|k) = s(k), && \text{(16d)}\\
& \forall i \in \{0, \ldots, N_p - 1\},
\end{aligned}$$

where $\mathbf{s}(k) = [s(0|k)^\top, \ldots, s(N_p|k))]^\top$ is the vector of predicted states $s(i|k)$ over the prediction horizon $N_p = 5$ at time k; $\mathbf{a}(k) = [a(0|k), \ldots, a(N_p - 1|k))]$ is the vector of predicted inputs $a(i|k)$ at time k; CEM_{sp} is the setpoint for the CEM; $\mathcal{S} = [25°\mathrm{C}, 0$ arb. units$] \times [45°\mathrm{C}, 80$ arb. units$]$ is the set of state constraints; $\mathcal{A} = [1.5$ W$, 1.5$ SLM$] \times [5$ W$, 5$ SLM$]$ is the set of input constraints; $h_c(s,a) = \hat{A}s + \hat{B}a$ is the control-relevant state space model that may differ from (14); and $\mathrm{CEM}_c(N_p|k) = \mathrm{CEM}(k) + \sum_{i=0}^{N_p-1} \hat{K}^{T_{\mathrm{ref}} - T(i|k)}\delta t$ is the control-relevant CEM model that may differ from (15). The solution to (16) defines the MPC law as

$$\pi_c\left(s(k)\right) = a^\star(0|k), \tag{17}$$

where $a^\star(0|k)$ is the optimal first input. The MPC problem (16) is implemented in Python using CasADi [18] and is solved using IPOPT [19].

B. MPC Law Adaptation using SEBO

The plant (14) and (15) has parameters (A, B, K) that are specific to a given patient, which are not known in advance. The control method in (16), on the other hand, only has estimates of these values based on general population data. Therefore, to personalize a CAPJ treatment, we propose to use SEBO to adapt a subset of these parameters to improve closed-loop performance while remaining safe using as few iterations as possible. This problem can be formulated in terms of (1) by defining the following black-box functions

$$f^0(x) = \textstyle\sum_{k=0}^{N} (\mathrm{CEM}_{sp} - \mathrm{CEM}(k))^2, \tag{18a}$$
$$f^1(x) = \textstyle\sum_{k=0}^{N} ([T(k) - T_{\max}]^+)^2, \tag{18b}$$

where $x = [\hat{A}_{11}, \hat{A}_{12}, \hat{A}_{21}, \hat{A}_{22}, \hat{K}]^\top$ are the subset of model parameters in the controller that we allow to be modified (mainly the elements of the estimated A matrix and K constant since we assume B can be accurately identified), N denotes the treatment time, and $T_{\max}$ denotes the maximum allowed safe surface temperature. The objective (18a) corresponds to the cumulative deviation of the CEM value from its setpoint while the constraint (18b) corresponds to the squared summation of all constraint violations over the full course of the treatment. The CEM and temperature values in (18a) and (18b) correspond to the true closed-loop values obtained by applying the MPC law (17) to the plant (14)–(15). As such, every evaluation of these functions requires an expensive closed-loop experiment. We allow all model parameters to vary within the following geometric bounds,

$$\underline{x} = \begin{cases} \hat{x}_0/v, & \hat{x}_0 > 0, \\ v\hat{x}_0, & \hat{x}_0 \leq 0, \end{cases}, \quad \overline{x} = \begin{cases} v\hat{x}_0, & \hat{x}_0 > 0, \\ \hat{x}_0/v, & \hat{x}_0 \leq 0, \end{cases},$$

where $\hat{x}_0$ is the nominal value of x, $\underline{x}$ is the lower bound, $\overline{x}$ is the upper bound, and $v = 2$. These bounds define the search space $\mathcal{X}$.

C. Results

To demonstrate SEBO, we consider the CAP treatment of a subject over a time period of $N = 120$ s. The objective is to deliver $\mathrm{CEM}_{sp} = 1.5$ min of thermal dose as quickly as possible, while the constraint with $T_{\max} = 45°\mathrm{C}$ ensures safety and comfort of the individual subject. We consider a limited budget of 30 iterations for SEBO, where the objective (18a) and constraint (18b) are observed after each full treatment time of 120 s and are modeled as independent

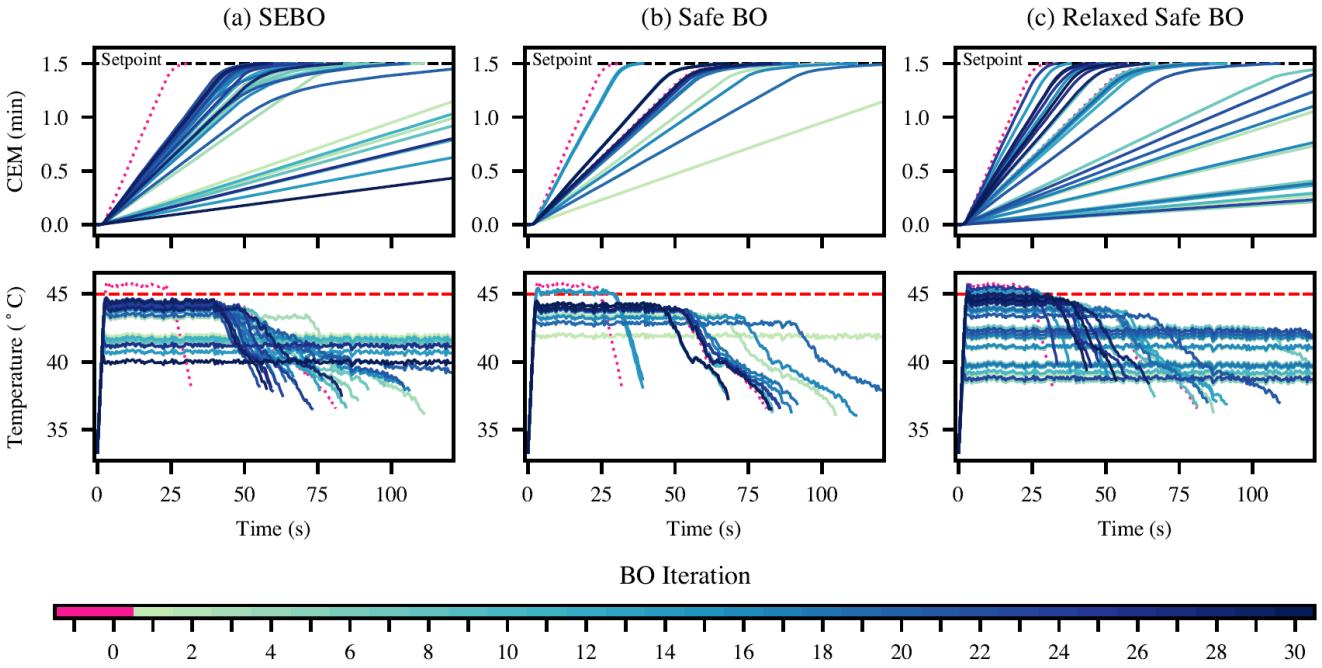

Fig. 2. Comparison of observed closed-loop profiles between three strategies: (a) SEBO, (b) safe BO, and (c) the relaxed formulation of safe BO. The top figures represent the evolution of CEM over a treatment period of 120 s. The bottom figures represent the evolution of temperature over the same treatment period. The colors/gradient of the profiles indicate the evolution of the profiles over 30 iterations of BO. The first two profiles in dotted pink indicate the initial data provided to BO.

GPs.[1] We select $\tau = 10^3$ and $\{\tilde{\beta}^i\}_{i=0}^m = 0.1$ as "poor" selection of the parameters from the standard safe BO, while $\varepsilon = 0.5$ and $b = 10^{-2}$ in Algorithm 1.[2] ε is chosen based on a user preference to indicate the level of desired potential improvement. b is chosen to be a small number near the neighborhood of projected query x_{n+1}^p. In this work, the sets $\hat{\mathcal{F}}_n$ and $\partial\hat{\mathcal{F}}_n$ are approximated using random samples.

Observed closed-loop trajectories of the CAP treatment for three BO methods are shown in Fig. 2. The three methods are compared column-wise: (a) SEBO, (b) standard safe BO (Section II), and (c) relaxed safe BO (BO using only (10)). The initial dataset $\mathcal{D}_0$ consists of one known feasible and one known unfeasible point in $\mathcal{X}$, shown in dotted pink. An infeasible initial data point helps to initially delineate between a safe region and unsafe region. In practice, such data points are readily estimated. For example, a controller that operates with low power at all times will most certainly provide a feasible solution, as power is directly related to surface temperature. Meanwhile, a controller that operates with high power at all times will most certainly provide an infeasible solution. The remaining profiles evolve over 30 iterations. Yellow-green profiles indicate earlier observations, and dark blue profiles indicate later observations. The top figures are the CEM profiles, which represent the objective (18a), and the bottom figures are the temperature profiles, which represent the constraint (18b). Looking at the profiles in Fig. 2(b), most of the search is contained near the initial feasible point, with very few exploratory actions. As a result, many iterations are not helpful in finding a better treatment. Meanwhile, the profiles in Fig. 2(c) demonstrate much more exploration of the design space $\mathcal{X}$, but with several constraint-violating queries. Using SEBO allows a middle-ground result. In Fig. 2(a), the observations exhibit more exploration compared to the standard safe BO, while still strictly adhering to the constraint. This result shows that SEBO uses and retains the information from the relaxed problem. We note that the standard safe BO appears to violate the constraint for two observations, and may be a result of a poor choice in β since the safety guarantees are probabilistic in nature.

Furthermore, we show that the revealed feasible set for SEBO is larger than what is estimated by standard safe BO in Fig. 3. By the end of 30 iterations, the posterior models can be used to visualize the revealed feasible set $\hat{\mathcal{F}}_{30}$. In Fig. 3, blue stars indicate $\hat{\mathcal{F}}_{30}$ of SEBO; orange circles indicate $\hat{\mathcal{F}}_{30}$ of standard safe BO; and green triangles indicate $\hat{\mathcal{F}}_{30}$ of the relaxed safe BO. We select a subset of parameters to examine and visualize, namely A_{11}, A_{22}, and K, due to their influence on the states s; A_{11} and A_{22} are the diagonal elements of A, and K is the exponential base of the CEM delivery. The subplots of Fig. 3 represent the 3 planes of the 3-dimensional space. Relaxed safe BO and SEBO both have larger $\hat{\mathcal{F}}_{30}$. In general, since the main influence of constraint violation involves temperature, A_{11} exhibits the most restrictive range of feasible parameters. Meanwhile, the influence of A_{22} and

[1]To implement SEBO, we modified components of Ax [20]. Ax interfaces with BoTorch [21] to perform BO. These tools were used with their default settings. Modifications to Ax are detailed in the codes available at `https://github.com/kchan45/SafeBOPlasma`.

[2]We note that $\tilde{\beta}$ is selected by the user, and $\beta = \tilde{\beta}\frac{\pi}{2}$ to match the implementation in Ax.

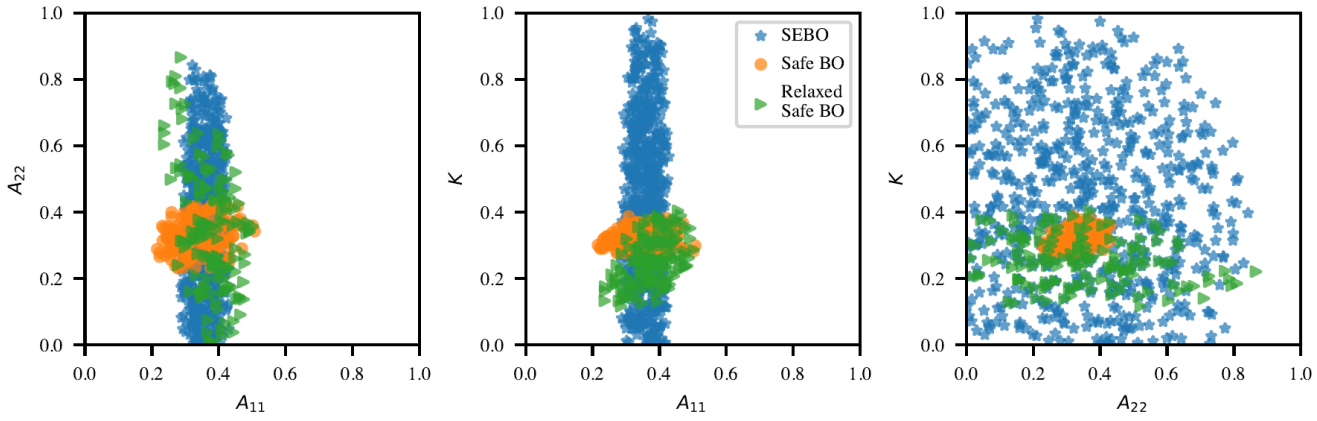

Fig. 3. Comparison of revealed feasible sets of SEBO, safe BO, and the relaxed formulation of safe BO at iteration 30; the feasible set is denoted by $\hat{\mathcal{F}}_{30}$. Blue stars indicate $\hat{\mathcal{F}}_{30}$ of SEBO; orange circles indicate $\hat{\mathcal{F}}_{30}$ of standard safe BO; and green triangles indicate $\hat{\mathcal{F}}_{30}$ of the relaxed safe BO. While 5 parameters were included in the design space of BO, i.e., $x \in \mathbb{R}^5$, we show the revealed set for 3 of the parameters A_{11}, A_{22}, and K since they are deemed the most influential to the objective and constraints. Values of the parameters are normalized.

K are less influential with respect to safety. A_{22} relates to the total optical intensity, and K describes the rate of thermal dose delivery (CEM), both of which have no influence on safety. Nonetheless, standard safe BO provides a myopic view of the design space. $\hat{\mathcal{F}}_{30}$ of relaxed safe BO has a more holistic view of what is deemed feasible/unfeasible, so its region is larger than standard safe BO, but smaller than SEBO because it has explored more areas of infeasibility.

V. Conclusions and Future Work

This paper presented a new method for safe explorative Bayesian optimization (SEBO) to improve the performance of standard safe BO that relies on penalizing constraint violations. SEBO ensures that the recommended designs do not get stuck in a locally feasible space. We demonstrated SEBO for run-to-run adaptation of an MPC law used to control cold atmospheric plasma treatment of surfaces in the context of plasma medicine. We showed that SEBO outperforms standard safe BO by exploring safe regions on the boundary in the direction of objective improvement. Future work will involve evaluating the volume of improvement of the feasible region in conjunction with performance improvement.

References

[1] T. Zhang, Q. Li, C.-s. Zhang, H.-w. Liang, P. Li, T.-m. Wang, S. Li, Y.-l. Zhu, and C. Wu, "Current trends in the development of intelligent unmanned autonomous systems," *Front. Inf. Technol. Electron. Eng.*, vol. 18, pp. 68–85, 2017.

[2] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. De Freitas, "Taking the human out of the loop: A review of Bayesian optimization," *Proc. IEEE*, vol. 104, no. 1, pp. 148–175, 2015.

[3] C. E. Rasmussen, C. K. Williams, *et al.*, *Gaussian processes for machine learning*, vol. 1. Springer, 2006.

[4] D. Krishnamoorthy and F. J. Doyle, "Safe Bayesian optimization using interior-point methods—applied to personalized insulin dose guidance," *IEEE Control Syst. Lett.*, vol. 6, pp. 2834–2839, 2022.

[5] F. Berkenkamp, A. Krause, and A. P. Schoellig, "Bayesian optimization with safety constraints: Safe and automatic parameter tuning in robotics," *Mach. Learn.*, pp. 1–35, 2021.

[6] D. Krishnamoorthy and F. J. Doyle III, "Model-free real-time optimization of process systems using safe Bayesian optimization," *AIChE J.*, vol. 69, no. 4, p. e17993, 2023.

[7] M. Laroussi, S. Bekeschus, M. Keidar, A. Bogaerts, A. Fridman, X. Lu, K. Ostrikov, M. Hori, K. Stapelmann, V. Miller, *et al.*, "Low-temperature plasma for biology, hygiene, and medicine: Perspective and roadmap," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 6, no. 2, pp. 127–157, 2022.

[8] A. D. Bonzanini, K. Shao, A. Stancampiano, D. B. Graves, and A. Mesbah, "Perspectives on machine learning-assisted plasma medicine: Toward automated plasma treatment," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 6, no. 1, pp. 16–32, 2021.

[9] D. Gidon, D. B. Graves, and A. Mesbah, "Predictive control of 2D spatial thermal dose delivery in atmospheric pressure plasma jets," *Plasma Sources Sci. Technol.*, vol. 28, no. 8, p. 085001, 2019.

[10] M. A. Hamburg and F. S. Collins, "The path to personalized medicine," *N. Engl. J. Med.*, vol. 363, no. 4, pp. 301–304, 2010.

[11] J.-M. Pouvesle and E. Robert, "Multimodal action of atmospheric pressure plasma jets for biological applications," in *ISPB 2017*, 2017.

[12] S. R. Chowdhury and A. Gopalan, "On kernelized multi-armed bandits," in *Proc. 34th Int. Conf. on Machine Learning*, pp. 844–853, PMLR, 2017.

[13] J. A. Paulson and C. Lu, "COBALT: COnstrained Bayesian optimizAtion of computationaLly expensive grey-box models exploiting derivaTive information," *Comput. Chem. Eng.*, vol. 160, p. 107700, 2022.

[14] C. Lu and J. A. Paulson, "No-regret Bayesian optimization with unknown equality and inequality constraints using exact penalty functions," *IFAC-PapersOnLine*, vol. 55, no. 7, pp. 895–902, 2022.

[15] P. Van Overschee and B. De Moor, *Subspace identification for linear systems: Theory—Implementation—Applications*. Springer Science & Business Media, 2012.

[16] D. Gidon, D. B. Graves, and A. Mesbah, "Effective dose delivery in atmospheric pressure plasma jets for plasma medicine: A model predictive control approach," *Plasma Sources Sci. Technol.*, vol. 26, no. 8, p. 085005, 2017.

[17] S. A. Sapareto and W. C. Dewey, "Thermal dose determination in cancer therapy," *Int. J. Radiat. Oncol. Bio. Phys.*, vol. 10, no. 6, pp. 787–800, 1984.

[18] J. A. Andersson, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl, "CasADi: a software framework for nonlinear optimization and optimal control," *Math. Program. Comput.*, vol. 11, pp. 1–36, 2019.

[19] A. Wächter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Math. Program.*, vol. 106, pp. 25–57, 2006.

[20] E. Bakshy, L. Dworkin, B. Karrer, K. Kashin, B. Letham, A. Murthy, and S. Singh, "AE: A domain-agnostic platform for adaptive experimentation," in *Proc. Conf. Neural Information Processing Systems*, 2018.

[21] M. Balandat, B. Karrer, D. Jiang, S. Daulton, B. Letham, A. G. Wilson, and E. Bakshy, "BoTorch: A framework for efficient Monte-Carlo Bayesian optimization," *Adv. Neural Inf. Process.*, vol. 33, pp. 21524–21538, 2020.