IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 6 November 2024, accepted 23 December 2024, date of publication 25 December 2024, date of current version 3 January 2025.

Digital Object Identifier 10.1109/ACCESS.2024.3522876

SURVEY

Constrained Bayesian Optimization: A Review

SASAN AMINI[1,2,3], **INNEKE VANNIEUWENHUYSE**[1,2,3], **AND ALEJANDRO MORALES-HERNÁNDEZ**[4]

[1]Flanders Make@UHasselt, Hasselt University, 3590 Diepenbeek, Belgium
[2]Data Science Institute, Hasselt University, 3590 Diepenbeek, Belgium
[3]Computational Mathematics Research Group, Faculty of Sciences, UHasselt, 3590 Diepenbeek, Belgium
[4]Machine Learning Group, Faculté des Sciences, Université Libre de Bruxelles, 1050 Brussels, Belgium

Corresponding author: Sasan Amini (sasan.amini@uhasselt.be)

This work was supported by the Flanders Artificial Intelligence Research Program (FAIR-2).

ABSTRACT Bayesian optimization is a sequential optimization method that is particularly well suited for problems with limited computational budgets involving expensive and non-convex black-box functions. Though it has been widely used to solve various optimization tasks, most of the literature has focused on unconstrained settings, while many real-world problems are characterized by constraints. This paper reviews the current literature on single-objective constrained Bayesian optimization, classifying it according to three main algorithmic aspects: (*i*) the metamodel, (*ii*) the acquisition function, and (*iii*) the identification procedure. We discuss the current methods in each of these categories and conclude by a discussion of real-world applications and highlighting the main shortcomings in the literature, providing some promising directions for future research.

INDEX TERMS Bayesian optimization, constrained optimization, expensive black-box functions, Gaussian processes.

I. INTRODUCTION

Bayesian optimization (BO) is a sequential optimization method that is particularly well suited for problems involving expensive, non-convex black box functions [1], [2], [3], [4]. In the engineering field, the approach is often used in process and/or product design optimization [5], [6], [7], where experiments might involve computationally intensive simulations (such as computational fluid dynamics (CFD), [8], [9]), or costly experiments [10]. By smartly designing experiments and refining models iteratively during the optimization process, BO succeeds in solving problems while limiting the number of experiments needed.

Bayesian optimization is based on the concept of *Bayesian inference* [11]: each observation obtained by querying the function modifies the model's belief about the function's behaviour and influences the selection of the next query location. It has been shown that BO is very powerful for solving real-world problems with limited evaluation budgets (see e.g., [6], [12], [13]), as it can automatically balance exploration and exploitation, resulting in efficient use of the evaluation budget.

The majority of the BO literature so far has focused on single-objective and unconstrained settings, meaning that the optimum can be anywhere in the search space (which is typically a hyperrectangle). Nevertheless, in recent years, constrained BO (CBO) has gained increasing attention [1], [2], [14], [15], [16] where specific output constraints need to be met in this type of problem. Imposing constraints onto the optimization problem restricts the feasible solution space in an often complex, even uncertain manner [1]. This paper surveys the approaches currently available for single-objective CBO and categorizes these based on algorithmic aspects (the metamodel used, the acquisition function, and the procedure applied to identify the optimal solution). In general, we consider approaches aimed at solving single-objective constrained optimization problems of the following type:

$$\begin{aligned} &\min \mathbf{x} \in D f(\mathbf{x}, \xi_{\mathbf{x}}) \\ &\text{s.t.} \quad c_q(\mathbf{x}, \xi_{\mathbf{x}}) \leq 0, \; q \in \mathbb{Z}^* \end{aligned} \tag{1}$$

where $\mathbf{x}$ is a d dimensional input vector, $D = \{\mathbf{x} \in \mathbb{R}^d : \mathbf{l} \leq \mathbf{x} \leq \mathbf{u}\}$ defines the hyperrectangle limited by the lower

The associate editor coordinating the review of this manuscript and approving it for publication was Emanuele Crisostomi.

bound vector $\mathbf{l}$ and the upper bound vector $\mathbf{u}$, $f : D \rightarrow \mathbb{R}$ denotes a scalar-valued objective function, and each $c_q : D \rightarrow \mathbb{R}$ denotes a scalar-valued constraint function (Q in total). Here, $\mathbb{Z}^*$ represents the set of positive integers. The objective and the constraint functions are expensive black-box functions, that can only be evaluated through (computer or physical) experiments. Each of these function evaluations may be affected by noise (denoted by $\xi_{\mathbf{x}}$), for instance, because of inherent randomness or measurement errors. This noise may have constant variance across the search space (referred to as *homogeneous noise*, which is independent of the location $\mathbf{x}$), or may differ across input locations (*heterogeneous noise*). In the absence of noise, the function evaluations are deterministic and reflect the true function - i.e., $\mathbb{E}\left[f(\mathbf{x}, \xi_{\mathbf{x}})\right] = f(\mathbf{x})$ and $\mathbb{E}\left[c_q(\mathbf{x}, \xi_{\mathbf{x}})\right] = c_q(\mathbf{x})$.

The *feasible region* of such an optimization problem consists of the area of the search space D where all constraints are satisfied. Note that this does *not* imply that evaluations of the goal and constraint functions outside the feasible region are by definition useless. On the contrary, in CBO, the algorithm is often allowed to deliberately sample in the infeasible region to gain information about the feasible region borders [17]. In that respect, CBO differs from Safe BO [14]: the latter does not allow (at least with a given probability) any function evaluation outside the feasible region, as this could lead to critical system failures or could destroy the system [18]. For readers interested in Safe BO, we refer to [18] and [19] for more details.

We focus on the CBO literature published between 2000 and 2024 in peer-reviewed journals, conference proceedings, and book chapters. We performed a WoS (Web of Science) search, using keywords such as "constrained Bayesian optimization", "constrained black-box optimization", and "constrained simulation optimization". By applying the ancestry approach [20], we selected a set of 48 articles (as shown in Table 1) that are relevant for this review. To the best of our knowledge, our work presents the first comprehensive survey for single-objective *constrained* BO algorithms with a specific focus on algorithmic aspects. Previous reviews on Bayesian optimization [1], [14], [21], [22], [23] have focused mainly on unconstrained algorithms; CBO algorithms were at best discussed on the sidelines, listing them as "exotic" or as "BO extensions". A very brief overview of expensive constrained black-box optimization has been provided in [24]. The study presented in [25] focuses on the characteristics of the constraints (which are not necessarily expensive) in simulation optimization problems. They provide definitions for several constraint classes, yet without discussing any algorithmic aspect. In the current work, we focus explicitly on settings with expensive black-box functions, and pay specific attention to the algorithmic design.

The remainder of this article is organized as follows. Section II introduces the topic of constrained Bayesian optimization, identifying the key algorithmic aspects in constrained BO algorithms. Section III discusses the main findings of this review and the pros and cons of the different approaches. Finally, Section IV summarizes the findings, highlighting potential improvements and avenues for further research.

II. CONSTRAINED BAYESIAN OPTIMIZATION

Bayesian optimization facilitates *efficient* exploration and exploitation of complex black-box objective functions. It leverages the principles of *Bayesian inference* [11], which allows for the incorporation of prior knowledge and the continuous updating of beliefs based on sequentially observed data, providing a framework for estimating outputs of a system of interest [1], [26].

All BO algorithms share some common elements in their workflow, as visualized in Figure 1. The key aspects in which constrained BO algorithms differ from their unconstrained counterparts are highlighted in the figure and form the basis for a more detailed discussion of the selected papers in Section III.

A set of **initial design** points is generated in the first step. Space-filling designs are to be preferred, such that information is gathered (ideally) across the entire design space [23]. Various mathematical techniques and algorithms are employed to generate space-filling designs, such as Latin hypercube sampling [27] or Sobol sequences [28].

Next, the BO algorithm enters a loop. **Evaluations** (by means of an expensive simulator/physical experiment) provide information about the unknown function(s) at the selected design points. Based on this information, the belief about the functions will be iteratively updated [3], and new points to be evaluated are selected by means of an acquisition function. In deterministic settings, the function output is solely determined by the input parameters, so one evaluation at each design point suffices. In stochastic settings, the function's output is subject to randomness or uncertainty. Multiple evaluations (referred to as *replications*) might then be required to gain a clearer insight into the estimated outcome (and variability on this outcome) at each point.

The **metamodel** is a key component in any BO algorithm, as it enables the algorithm to estimate the function's behaviour at unobserved design points, and to guide the optimization process [3], [14], [29]. Gaussian processes (GP) (also known as *kriging* models) are the most popular type of metamodel in BO algorithms, as they provide a natural way to quantify the uncertainty of the model outcome at unobserved design points, which facilitates the efficient exploration/exploitation of the search space [2], [30], [31].

In the most basic GP formulation (see e.g., [27]), the underlying function is assumed to be deterministic. Yet, many real-world settings involve noise [5], [15]. A GP with nugget effect [14] can be used to handle noisy observations that are caused by measurement errors or small-scale variability [16], [32], [33]. In this model, the noise variance is assumed to be homogeneous across the search space. The nugget effect is added to the GP model and is reflected in the resulting predictor uncertainty. When the underlying

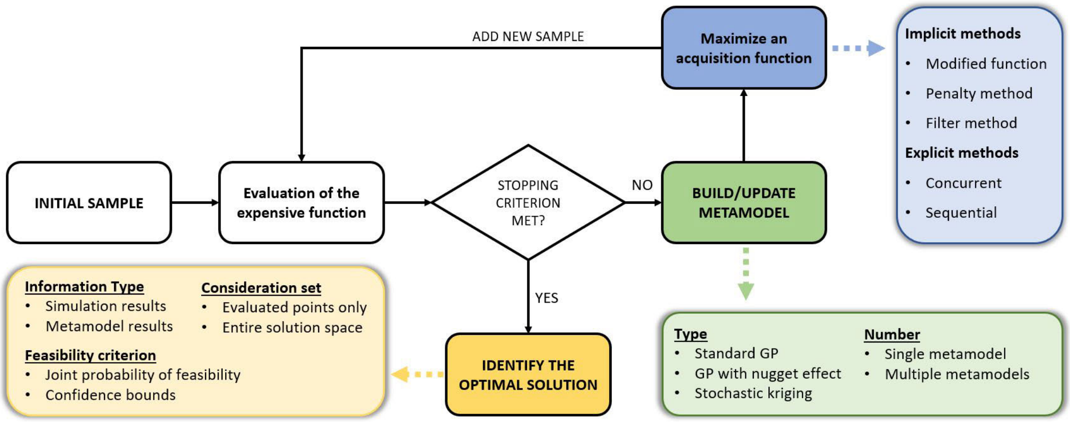

FIGURE 1. General workflow of Bayesian optimization algorithms, highlighting the key steps in which constrained algorithms differ from their unconstrained counterparts.

function is inherently random, the noise variance is usually heterogenous. Specific GP approaches exist to approximate the functions in such settings. *Stochastic kriging* [35], for instance, estimates the intrinsic uncertainty (inherent in the stochastic system) at the different observed points, which impacts the resulting values for the predictor and predictor uncertainty at both observed and non-observed points.

In constrained Bayesian optimization, multiple unknown functions are naturally involved (objective and constraint functions), so the most natural approach is to fit multiple metamodels, each one corresponding to an individual unknown function. Yet, some algorithms fit one single metamodel to an augmented function that encompasses both objective and constraint functions, as further detailed in Section III-A.

The **acquisition function** (also referred to as *infill criterion* [35]) is the second key component of any BO algorithm [1]. It uses the metamodel information (i.e., the probabilistic belief of the function's behavior) to quantify the attractiveness of new solutions to be evaluated by the expensive simulator [1]. The point where the acquisition function is maximized then is the point that is sampled next. In unconstrained BO, the acquisition function is typically designed to automatically balance exploitation (i.e., sampling in areas with promising predictor values) and exploration (sampling in areas with high uncertainty on the objective function). Popular acquisition functions in unconstrained BO are Probability of Improvement (PI, [36]) and Expected Improvement (EI, [27]). For further insights into various versions of EI-based acquisition functions, we refer to [37]. Recently, there has been a growing interest in entropy-related acquisition functions, which focus on reducing the uncertainty about the location of the optimum [1].

The presence of constraints in CBO leads to changes in the sampling procedure. Following [33] and [38], we distinguish two main methods. In the *implicit method*, a new acquisition function is defined to incorporate the effect of constraints using a merit-type function. In the *explicit method*, by contrast, an *un*constrained acquisition function is employed, in each iteration, to evaluate the attractiveness of the points *within* the estimated feasible region. Further details are discussed in Section III-B.

The BO algorithm is **terminated** when a stopping condition is met; e.g., when the algorithm has reached a satisfactory solution [39], or when the available budget has been depleted [40]. Then, the algorithm needs to **identify the optimal solution(s)** in the final step. In the current BO (and CBO) literature, this last step is often considered to be trivial. Yet, clear differences in approach can be identified, depending on three aspects: the *information* used (either simulation results or metamodel results), the *consideration set* (which may be limited to the evaluated points only or may consider both evaluated and unobserved solutions), and the *feasibility criterion* (the joint probability of feasibility (PoF), or individual confidence bounds (CB)). We discuss this further in Section III-C.

Bayesian Optimization is highly effective for optimizing expensive, black-box functions but faces several computational challenges. The key issue is the common reliance on GPs, which scale poorly with data size due to their $O(n^3)$ complexity for training and $O(n^2)$ for storage, with n as the number of observed design points. In CBO, both the objective and constraint functions are typically modelled using GPs to account for uncertainty and to limit the number of expensive evaluations. This adds to the increase in the computational complexity of unconstrained BO, especially when the number of constraints grows. The key challenge in CBO lies in balancing exploration (searching for feasible regions) and exploitation (optimizing within feasible areas) since the feasibility threshold is often unknown in the input space. In general, optimizing acquisition functions is non-trivial and often requires non-convex and computationally expensive procedures, especially in high-dimensional spaces.

In expensive black-box settings, the computational complexity of an optimization algorithm can be evaluated by means of the *worst-case expected running time* [41]. The running time (or *optimization time*) of an algorithm for a given function f can be measured by the*number of function evaluations* that the algorithm performs until (and including) the evaluation of an optimal solution for f. For constrained problems, the running time is largely proportional to the number of objective and constraint evaluations, as these evaluation times usually dominate the time required to perform the remainder of the algorithm's instructions.

III. RESULTS AND DISCUSSION

Table 1 offers an overview of the CBO studies we surveyed, classified according to the algorithmic aspects discussed above. In what follows, we discuss our findings, highlighting the advantages and limitations of the different approaches. Additionally, Table 1 offers an overview of real-world problems solved by these algorithms, highlighting the practical applications of CBO methods, which will be further discussed at the end of this section.

A. METAMODELS

Considering that the majority of the surveyed references focus on the deterministic setting, it is no surprise that the standard Gaussian Process is mostly considered as the metamodel type. In [15], it is used in a stochastic setting; the algorithm replicates the function outcome at each design point and then assumes that the obtained sample mean is a perfect estimator of the underlying function. This is a simple, yet naive approach; it is often used to reduce noise on output estimates in evolutionary approaches (see, e.g., [42]), yet in settings with a small budget and an expensive simulator, the opportunity for replications will likely be very limited. Indeed, replicating already observed inputs reduces the simulation budget for observing new points, which will impact algorithm performance [35]. Certainly in settings with substantial noise on the observations, the resulting sample means may then still exhibit a lot of uncertainty.

The few CBO papers that focus on noisy settings account for homoscedastic noise in the metamodel, using a GP with a nugget effect to estimate the unknown objective and constraint functions. However, many real-world problems are characterized by heteroscedastic noise; some well-known examples include the waiting time behaviour in the M/M/1 queueing system [43] and the travel time in a freeway traffic management system [44]. The algorithms proposed in [45] and [46] are the only exceptions, accounting for heteroscedastic noise through a stochastic kriging metamodel. Their results show significant improvements in the choice of sampled points compared to the use of a standard GP.

Constrained BO involves more than one unknown function; the objective function, and at least one constraint function. To deal with multiple unknown functions, some algorithms use an augmentation method to consider all function outputs simultaneously in a single objective function. Hence, a *single* metamodel can be fit to the resulting transformed function. The main benefit of this approach is that the closed-form expressions of many acquisition functions can be used [37]. Yet, the augmented function is often much more complex than the individual objective and constraint functions, which makes it more difficult for the metamodel to approximate [47]. In [48], the *particle learning multivariate Gaussian process* (PLMGP) [49] is employed to model correlated outputs of the simulation model without the use of an augmentation function.

In most articles, multiple independent metamodels are fit to the objective function and the individual constraints. The authors in [48] observed in their experiments that using multiple independent GPs yields results similar to fitting a single GP.

B. ACQUISITION FUNCTION

The existence of constraints necessitates adjustments to the search step and sampling procedure. We distinguish two main approaches in the CBO literature: *implicit methods* and *explicit methods*.

1) IMPLICIT METHODS

Algorithms in this class avoid explicitly solving a constrained problem. They define a merit-type acquisition function that incorporates the effect of constraints in the sampling policy [38]. Many implicit acquisition functions are modified versions of unconstrained acquisition functions to which a feasibility element has been added.

Constrained Expected Improvement (CEI) [15], [32], for instance, is widely used. It consists of the well-known EI function [27] for unconstrained optimization, multiplied by a factor that estimates the *joint probability of feasibility (PoF)* of the solution:

$$CEI(\mathbf{x}) = EI(\mathbf{x}) \times PoF(\mathbf{x}) \tag{2}$$

$$PoF(\mathbf{x}) = \prod_{q=1}^{Q} \Big[Pr\Big(c_q(\mathbf{x}) \leq 0\Big) \Big] \tag{3}$$

The notation $Pr(.)$ reflects the probability that constraint q is satisfied. When a GP is used to model the constraint function, $Pr\Big(c_q(\mathbf{x}) \leq 0\Big) = \Phi(\frac{-\hat{\mu}_q(\mathbf{x})}{\hat{\sigma}_q(\mathbf{x})})$ at any point $\mathbf{x}$, where $\hat{\mu}_q(\mathbf{x})$ and $\hat{\sigma}_q(\mathbf{x})$ refer to the predictor and the predictor uncertainty for constraint q at the point $\mathbf{x}$, respectively and Φ denotes the cumulative normal probability density function. Note that individual (independent) metamodels are required to calculate the Probability of Feasibility (PoF). Eq. (3) also assumes that the constraints are mutually independent [32], [50], [51]. Recent results in [52] indicate that incorporating dependence between the constraints does not significantly improve the performance of the algorithms.

The authors of [53] replace the EI criterion in Eq. (2) by the Probability of Improvement (PI), resulting in the Constrained Probability of Improvement (CPI). The CPI suffers from the

TABLE 1. Overview of constrained Bayesian optimization algorithms, classified according to their algorithmic aspects.

Reference	Problem setting	Metamodel		Acquisition function	Identification†			Real world application
		Type	Number		Con. Set	Info.	Feas.	
Sasena et al., 2002 [56]	Deterministic	Standard GP	Multiple	Implicit (modified)	NM	NM	NM	Hybrid electric vehicle (HEV) design problem (3d problem with 8 constraints)
Gramacy and Lee, 2011 [17]	Deterministic	Standard GP	Multiple	Implicit (penalty)	Complete	Metamodel	PoF	Health care policy optimization (6d problem with 2 constraints)
Lee et al., 2011 [57]	Deterministic	Standard GP	Multiple	Implicit (modified)	NM	NM	NM	Hydraulic capture problem (12d problem with 8 constraints)
Parr et al., 2012 [58]	Deterministic	Standard GP	Multiple	Implicit (filter)	NM	NM	NM	Satellite boom design problem (8d problem with 1 constraint), Aircraft wing box design problem (6d problem with 2 constraints)
Gardner et al., 2014 [50]	Deterministic	Standard GP	Multiple	Implicit (modified)	NM	NM	NM	Hyperparameter optimization (2 and 3 d problems with 1 constraint)
Gelbart et al., 2014 [32]	Noisy	GP with nugget	Multiple	Implicit (modified)	NM	NM	NM	Hyperparameter optimization (11d problem with 1 constraint)
Picheny et al., 2014 [59]	Deterministic	Standard GP	Multiple	Implicit (modified)	NM	NM	NM	Only synthetic functions
Lindeberg and Lee, 2015 [60]	Deterministic	Standard GP	Multiple	Implicit (modified)	NM	NM	NM	Hydraulic capture problem (3d problem with 1 constraints)
Durantin et al., 2016 [61]	Deterministic	Standard GP	Multiple	Implicit (filter)	NM	NM	NM	Solid propulsion design problem (4d problem with 3 constraints)
Gramacy et al., 2016 [47]	Deterministic	Standard GP	Multiple	Implicit (modified)	NM	NM	NM	Hydraulic capture problem (6d problem with 2 constraints)
Hernandez-Lobato et al., 2016 [16]	Noisy	GP with nugget	Multiple	Implicit (modified)	Complete	Metamodel	PoF	Hyperparameter optimization (12d problem with 1 constraint)
Picheny et al., 2016 [62]	Deterministic	Standard GP	Multiple	Implicit (penalty)	NM	NM	NM	Only synthetic functions
Bagheri et al., 2017 [63]	Deterministic	Standard GP	Multiple	Implicit (modified)	NM	NM	NM	Aerodynamic shape design problem (4d problem with 5 constraints)
Lam and Willcox, 2017 [64]	Deterministic	Standard GP	Multiple	Implicit (modified)	Complete	Metamodel	PoF	Only synthetic functions
Li et al., 2017 [65]	Deterministic	Standard GP	Multiple	Explicit (concurrent)	Eval. only	Simulation	Trivial	Spring design problem (3d problems with 4 constraints), Welded beam design problem (4d problem with 6 constraints), and Speed reducer design problem (7d problem with 11 constraints)
Carpio et al., 2018 [53]	Deterministic	Standard GP	Multiple	Implicit (modified)	Eval. only	Simulation	Trivial	Reactor network design (3d problem with 4 constraints),
Chung et al., 2018 [66]	Deterministic	Standard GP	Multiple	Implicit (penalty)	NM	NM	NM	Flange shaft design problem (7d problem with 1 constraint)
Dong et al., 2018 [67]	Deterministic	Standard GP	Multiple	Explicit (concurrent)	NM	NM	NM	Spring design problem (3d problem with 4 constraints), Welded beam design problem (4d problem with 7 constraints), Pressure vessel design problem (4d problem with 4 constraints), Speed reducer design problem (7d problem with 4 constraints), Stepped cantilever beam design (10d problem with 11 constraints)
Yuan et al., 2018 [68]	Deterministic	Standard GP	Single	Implicit (modified)	Eval. only	Simulation	Trivial	Satellite Design Problem (12 d problem with 7 constraints)
Wang and Ierapetritou, 2018 [46]	Noisy	Stochastic kriging	multiple	Implicit (modified)	Eval. only	Metamodel	PoF	Synthetic functions only
Ariafar et al., 2019 [69]	Deterministic	Standard GP	Multiple	Implicit (penalty)	Eval. only	Metamodel	PoF	Hyperparameter optimization (12d problem with 1 constraint)
Bartoli et al., 2019 [70]	Deterministic	Standard GP	Multiple	Explicit (concurrent)	NM	NM	NM	Wing design problem (17d problem with 1 constraint)
Jiao et al., 2019 [71]	Deterministic	Standard GP	Multiple	Implicit (modified)	Eval. only	Simulation	Trivial	Only synthetic functions
Letham et al., 2019 [15]	Noisy	Standard GP	Multiple	Implicit (modified)	Eval. only	Metamodel	PoF	Hyperparameter optimization (6d problem with 1 constraint), Server performance problem (7d problem with 1 constraint)
Shi et al., 2019 [72]	Deterministic	Standard GP	Multiple	Implicit (filter)	NM	NM	NM	Satellite design problem (15d problem with 10 constraints), Internal combustion engine design problem (5d problem with 9 constraints), I-beam design problem (4d problem with 2 constraints)
Tran et al., 2019 [73]	Noisy	GP with nugget	Multiple	Implicit (modified)	NM	NM	NM	Slurry pump impeller design problem (33d problem with 4 constraints)
Akbari and Kazerooni, 2020 [74]	Deterministic	Standard GP	Multiple	Implicit (modified)	Eval. only	Simulation	Trivial	10 different engineering design problems (2 to 10 d problems with 2 to 11 constraints)
Parnianifard, 2020 [75]	Deterministic	Standard GP	Single	Implicit (penalty)	Eval. only	Simulation	Trivial	Spring design problem (3d problem with 4 constraints), Welded beam design problem (4d problem with 7 constraints), Pressure vessel design problem (4d problem with 4 constraints)
Pourmohamad and Lee, 2020 [48]	Deterministic	PLMGP ‡	Single	Implicit (filter)	NM	NM	NM	Welded beam design problem (4d problem with 6 constraints), Hydraulic capture problem (6d problem with 2 constraints)
Priem, 2020 [38]	Deterministic	Standard GP	Multiple	Explicit (concurrent)	Complete	Metamodel	CB	Hybrid aircraft design (12d problem with 3 constraints)
Tao et al., 2020 [76]	Deterministic	Standard GP	Multiple	Explicit (sequential)	Complete	Metamodel	PoF	8 different engineering design problems (4 to 7 d problems with 2 to 11 constraints)
Jiang et al., 2021 [77]	Deterministic	Standard GP	Multiple	Implicit (filter)	Eval. only	Simulation	Trivial	I-beam design (4d problem with 2 constraints), Pressure vessel design problem (4d problem with 4 constraints), Speed reducer design problem (7d problem with 11 constraints)

TABLE 1. ***(Continued.)*** **Overview of constrained Bayesian optimization algorithms, classified according to their algorithmic aspects.**

Saves et al., 2021 [78]	Noisy	GP with nugget	Multiple	Implicit (modified)	NM	NM	NM	Aircraft design problem ($12d$ problem with 2 constraints), and Hybrid aircraft design problem ($12d$ problem with 4 constraints)
Siefman et al., 2021 [79]	Noisy	GP with nugget	Multiple	Implicit (modified)	NM	NM	NM	Nuclear waste tank design problem ($2d$ problem with 2 constraints)
Zhang et al., 2021 [80]	Deterministic	Standard GP	Multiple	Implicit (modified)	NM	NM	NM	Portfolio optimization problem ($8d$ problem with 2 constraints), and robot pushing problem ($3d$ problem with 1 constraint)
Zhang et al., 2021 [81]	Noisy	GP with nugget	Multiple	Implicit (filter)	Eval. only	Metamodel	PoF	Operational amplifier circuit design problem ($10d$ problem with 2constraints), Class-E power amplifier design problem ($12d$ problem with 1 constraint), Low-Power amplifier circuit design problem ($24d$ problem with 6 constraints), and Charge pump circuit design problem ($36d$ problem with 5 constraints)
Lu and Paulson, 2022 [33]	Noisy	GP with nugget	Multiple	Implicit (penalty)	Eval. only	Metamodel	CB	Synthetic functions only
Pourmohamad and Lee, 2022 [40]	Deterministic	Standard GP	Multiple	Implicit (penalty)	NM	NM	NM	Hydraulic capture problem ($6d$ problem with 2 constraints)
Zeng et al., 2022 [82]	Deterministic	Standard GP	Multiple	Explicit (sequential)	Eval. only	Simulation	Trivial	26 different engineering design problems (2 to 10 d problems with 2 to 12 constraints)
Amini and VanNieuwenhuyse, 2023 [45]	Noisy	Stochastic Kriging	Multiple	Implicit (penalty)	Complete	Metamodel	PoF	Synthetic function only
Nguyen et al., 2023 [83]	Deterministic	Standard GP	Multiple	Explicit (concurrent)	NM	NM	NM	Gas transmission compressor design problem ($4d$ problem with 1 constraint), Hyperparameter optimization problem ($5d$ problem with 10 constraints), and Quantum chip problem ($11d$ problem with 2 constraints)
Xu et al., 2023 [84]	Noisy	GP with nugget	Multiple	Explicit (concurrent)	NM	NM	NM	Temperature controller tuning problem ($3d$ problem with 1 constraint)
Bian et al., 2024 [85]	Noisy	GP with nugget	Multiple	Explicit (concurrent)	Eval. only	Metamodel	PoF	Airfoil design problem ($10d$ problem with 2 constraints)
Pelamatti et al., 2024 [86]	Deterministic	Multi-output GP	Multiple	Implicit (modified)	Eval. only	Metamodel	PoF	Compressor rotor design problem ($20d$ problem with 5 constraints)
Ragueneau et al., 2024 [87]	Deterministic	Standard GP	Multiple	Implicit (modified)	NM	NM	NM	Gantry crane design problem ($2d$ problem with 1 constraint)
Song et al., 2024 [88]	Deterministic	Standard GP	Multiple	Explicit (concurrent)	Eval. only	Metamodel	PoF	Ten-bar frame structure design problem ($15d$ problem with 1 constraint), Journal bearing lubrication problem ($6d$ problem with 1 constraint), and Water distribution network problem ($17d$ problem with 1 constraint)
Tfaily et al., 2024 [89]	Deterministic	Standard GP	Multiple	Implicit (modified)	Eval. only	Simulation	Trivial	Aircraft design problem ($12d$ problem with 12 constraints)
Ungredda and Branke, 2024 [51]	Noisy	GP with nugget	Multiple	Implicit (modified)	Eval. only	Metamodel	PoF	Hyperparameter optimization ($9d$ problem with 1 constraint)

† NM: not mentioned in the article
‡ PLMGP: Particle Learning Multivariate GP

same shortcomings as PI does, as it fails to quantify the extent of "*expected*" improvement in the objective function.

As an alternative, some authors (see e.g., [47], [48], [49], [54]) have opted to incorporate the impact of constraints employing approaches that are also used in the mathematical programming field, such as the penalty method [55]. In its classical form, this method transforms the constrained optimization problem in Eq. 1 into an unconstrained problem, by revising the objective function as follows:

$$\min_{\mathbf{x}\in D} \quad \mathcal{P}(\mathbf{x}, \xi_{\mathbf{x}})$$
$$\text{where} \quad \mathcal{P}(\mathbf{x}, \xi_{\mathbf{x}}) = f(\mathbf{x}, \xi_{\mathbf{x}}) + \sum_{q=1}^{Q} \left[M_q \cdot h\left(c_q(\mathbf{x}, \xi_{\mathbf{x}})\right) \right] \quad (4)$$

where $h\left(c_q(\mathbf{x}, \xi_{\mathbf{x}})\right)$ is defined as $max\left[0, c_q(\mathbf{x}, \xi_{\mathbf{x}})\right]$, and the M_q values are positive parameters designed to reduce the attractiveness of solutions that violate the constraints of the original problem. This ensures that the optimum of the revised objective function will be found within the original problem's feasible region. The resulting augmented function is then used directly as an acquisition function [69]. Alternatively, an unconstrained acquisition function such as EI has been employed [68].

Despite its apparent simplicity, this approach is non-trivial to implement, as the penalty parameter should be chosen such that feasibility is achieved without jeopardizing convergence to the true optimum. This can be a challenging task [1], [40], [90]. A low penalty value might result in the algorithm converging to a solution that significantly violates constraints, while a high penalty value could lead to suboptimal solutions or slow convergence. Selecting the right penalty value often necessitates a trial-and-error approach [14].

The algorithm proposed in [75] uses the death penalty method [90]: $M = \infty$. The acquisition function then exhibits a discontinuity at the estimated constraint boundary. This can pose challenges for optimization algorithms. For instance, algorithms that rely on the smoothness properties of the objective function to guide the search process will have difficulties converging when the optimum is at (or close to) the boundary.

To avoid this issue, alternative approaches such as barrier methods are used in the literature [45], [48]. Barrier methods, also known as *interior-point methods* [40], gradually increase the penalty as the search approaches the infeasible region. The barrier function used by [40] reformulates Problem (1) as:

$$\begin{aligned} &\min_{\mathbf{x}\in D} \quad \mathcal{B}(\mathbf{x}, \xi_{\mathbf{x}}) \\ &\text{where} \quad \mathcal{B}(\mathbf{x}, \xi_{\mathbf{x}}) = f(\mathbf{x}, \xi_{\mathbf{x}}) + b\sum_{q=1}^{Q} \log\big(-c_q(\mathbf{x}, \xi_{\mathbf{x}})\big) \end{aligned} \tag{5}$$

where b is a small positive constant. Each function (objective and constraints) is estimated by means of an independent GP, and Eq. (5) is used as the acquisition function. This barrier function will never sample solutions that are estimated to be infeasible, as $log\big(-c_q(\mathbf{x}, \xi_{\mathbf{x}})\big)$ is not defined when for $c_q(\mathbf{x}, \xi_{\mathbf{x}}) > 0$. It also penalizes points close to the border of the feasible region, and thus encourages sampling interior points. When the optimum point *is* an interior point, this might suffice. However, when at least one of the constraints is binding in the optimum, the performance of the algorithm is likely to suffer: as the acquisition function discourages sampling near the border, the metamodel estimates in that region will remain inaccurate. This may lead to the identification of a suboptimal solution as the final solution.

In problems with binding constraints, an appropriate b value improves the quality of the final solution returned. Theoretically, as b approaches zero, the minimum of $\mathcal{B}(\mathbf{x}, \xi_{\mathbf{x}})$ converges to the optimal solution of Problem (1). For this reason, the authors of [40] propose to replace b by the predictor uncertainty: $b(\mathbf{x}) = \hat{\sigma}_f^2(\mathbf{x})$. The value of $\hat{\sigma}_f^2(\mathbf{x})$ automatically decreases as more data points in a given area are sampled: consequently, the barrier penalty $b(\mathbf{x})$ decreases and we will automatically be able to push closer to the boundary. This approach avoids the introduction of an abrupt discontinuity in the goal function of Eq. (5) when $c_q(\mathbf{x}, \xi_{\mathbf{x}}) > 0$, and is also used by [45].

Another penalty approach known as the augmented Lagrangian method is used in [47] and [62]. With this method, Problem (1) is reformulated as:

$$\begin{aligned} &\min_{\mathbf{x}\in D} \quad \mathcal{L}(\mathbf{x}, \xi_{\mathbf{x}}) \\ &\text{where} \quad \mathcal{L}(\mathbf{x}, \xi_{\mathbf{x}}) = f(\mathbf{x}, \xi_{\mathbf{x}}) + \sum_{q=1}^{Q} \lambda_q \cdot c_q(\mathbf{x}, \xi_{\mathbf{x}}) \\ &\qquad\qquad + \frac{1}{2M}\sum_{q=1}^{Q} h\big(c_q(\mathbf{x}, \xi_{\mathbf{x}})\big)^2 \end{aligned} \tag{6}$$

where $h\big(c_q(\mathbf{x}, \xi_{\mathbf{x}})\big)$ is defined as $max\big[0, c_q(\mathbf{x}, \xi_{\mathbf{x}})\big]$, M is a positive penalty parameter, and λ_q ($q = 1, \ldots, Q$) are the Lagrange multipliers. This method does not introduce discontinuities and does not explicitly prevent infeasible solutions from being sampled. It iteratively optimizes the objective function in Eq. (6), updating the value of M and the Lagrange multipliers, to find a solution that simultaneously minimizes the objective function and satisfies the constraints. Interested readers can find more information in [47] and [62].

Last but not least, the *filter method* also provides an implicit way to solve CBO problems. This approach borrows ideas from the multiobjective optimization literature, and solves the constrained optimization problem in Equation 1 by trading off the attractiveness of points w.r.t. the objective function, and w.r.t. feasibility considerations. The aim is to identify the set of *non-dominated solutions*: i.e., those solutions for which no objective can be improved without worsening the other objective [91]. For example, [61] (and, analogously, [58]) use the probability of feasibility and the expected improvement as the two objectives to be maximized in their algorithm. The nondominated solutions are identified by means of the NSGA-II algorithm [92], and the point with the largest *CEI* value (see Eq. (2)) is then selected as the next point to evaluate. They point out that this point does not necessarily correspond to the point determined using the CEI acquisition function: the CEI function is severely multimodal, so treating the problem as multi-objective yields better solutions.

The authors of [61] also propose a three-objective approach, where the predictor uncertainty of the constraints is considered as the third objective to be minimized. The final point to be evaluated is again the non-dominated solution with the highest *CEI* value. The three-objective approach results in a lower optimality gap, yet this comes at the cost of more function calls. This is likely caused by the fact that the predictor uncertainty of the constraints (i.e., the third objective) encourages more exploration, also in areas that are not necessarily relevant from the perspective of feasibility or EI value. In constrained Bayesian optimization, the focus should primarily be on improving the accuracy of the constraint metamodels in areas where the objective function shows promising performance [93].

2) EXPLICIT METHODS

Algorithms using implicit methods solve an unconstrained problem, and can thus benefit from the progress made in the field of unconstrained optimization. Yet, the computational cost of these methods prohibits their use in large-scale optimization problems [38]. Explicit methods also use acquisition functions from the unconstrained BO literature to decide where to sample next. The difference is that, at each iteration, the feasible region is estimated based on the available data, and the acquisition function selects the next solution to sample *within* that region.

The most common approach is to estimate the feasible region using the GP predictions for the constraints [56]. However, this may lead to issues in the early iterations of the algorithm, when the metamodel is not yet sufficiently accurate [38]. Using confidence bounds helps to overcome this issue: for example, in [84], the individual lower confidence bounds of the constraints are used to identify the feasible region at each iteration. Then the solution with the minimum lower bound of the objective function is selected to query. The proposed algorithm in [38] takes a similar approach to declare the feasible region, but then uses the unconstrained acquisition function from [70] to select the point to evaluate.

The few papers using an explicit method mostly address the feasible region identification and the optimization problem by *concurrently* updating the information about the feasibility of the constraints and the behaviour of the goal function in each iteration. The only exceptions are the algorithms by [76] and [82]: these algorithms use a *sequential* approach, in which a first acquisition function is specifically used to identify the feasible region, while a second acquisition function is then used to select the next point within this estimated feasible region. In [76], for instance, the feasible region is identified using the predictor uncertainty of the constraints' metamodels as an acquisition function. Only when the desired accuracy level is reached for these metamodels, they locate the estimated optimum solution with a different acquisition function. Such a *sequential* approach avoids the challenge of simultaneously handling feasibility detection and optimization, yet it comes at a cost: the scarce computation budget may not be spent in the most data-efficient way, as samples in the first phase are allocated *without* considering the objective function. Here, the analyst is mainly interested in accurate estimations of the feasible region's borders in areas where the objective function has promising values. Moreover, further information on the constraint values may be gathered during the optimization phase, yet current algorithms do not use this information to modify their beliefs about the feasible region.

C. IDENTIFICATION OF THE BEST SOLUTION

The identification step has not received much attention so far in the CBO literature. In many papers, the authors do not even mention explicitly how they identify the final optimal solution (''**NM**'' in Table 1).

When studying a deterministic problem setting, and limiting the final consideration set to the evaluated solutions, identifying the best solutions will be *trivial*. For the other articles, we distinguished between three aspects: the ''information type'', the ''consideration set'', and the ''feasibility criterion'' (see Figure 2).

The information used to identify the optimum can either consist of simulation results or metamodel results. If simulation results are used, only the evaluated points can be considered in the identification step. In this survey, all articles using this approach studied deterministic settings. Consequently, there is no specific feasibility criterion required in the identification step, as the simulated values directly show whether a point is feasible.

When metamodel information is used, the consideration set is either limited to the evaluated points (e.g., [53] or [32]), or to the entire solution space (both evaluated and unobserved solutions). Each has its pros and cons. As there is no reason to believe that the best point will by definition be part of the evaluated set, and the metamodel information yields predictions (and prediction uncertainties) for the entire solution space, it may seem straightforward to not only consider evaluated points. On the other hand, relying on metamodel estimates at unobserved points may be overconfident, given that the model's belief of the function behavior will never coincide perfectly with the true function. Regardless of which consideration set is chosen, the criterion for checking feasibility is either the *joint probability of feasibility* (PoF, Equation 3), or a check on the individual confidence bounds of the constraint values. In the first case, all solutions that meet a user-defined minimum threshold for PoF are classified as feasible solutions. Only two algorithms, [33] and [38], use the individual confidence bounds of the constraints: a solution is then considered feasible when all individual upper confidence bounds (CB) are negative (for constraints of the form $c_q(\mathbf{x}, \xi_{\mathbf{x}}) \leq 0$, as shown in Problem 1).

D. DISCUSSION OF REAL-WORLD APPLICATIONS

Constrained Bayesian Optimization (CBO) has been applied to a variety of real-world applications, primarily in the domain of engineering design. These problems, including structural design [88], mechanical optimization [38], and circuit design [81], need to be optimized under constraints using expensive Computational Fluid Dynamics (CFD) simulation models, making CBO an ideal approach. While a significant portion of the research addresses these engineering problems, there is also growing interest in hyperparameter optimization in machine learning [51], an area where CBO methods have been successfully applied to optimize models under performance constraints.

Some attention has also been given to healthcare optimization problems [17], [94]; although this area remains relatively underexplored, it shows promise for future CBO applications. Surprisingly, operations management and logistics applications, areas that inherently deal with constrained optimization, are largely absent in the literature. This is likely due to the fact that many problem settings in these fields are combinatorial and high-dimensional, which presents significant challenges for CBO algorithms.

Most of the real-world applications reviewed involve less than 20 decision variables, reflecting the current limitations of CBO methods in handling high-dimensional problems. Exceptions are [73] and [81], who have considered high-dimensional problems (up to 35 decision variables). The curse of dimensionality is a well-known challenge in optimization [95], [96], with high-dimensional Bayesian optimization as a very active area in current BO research [96], [97].

Real-world problems often come with characteristics that inspire researchers to develop specific novel approaches in their algorithms. An example is the decoupled evaluation of objective and constraint functions [16], [83], [86], which allows for more flexible optimization processes. In drug discovery [98], for instance, researchers aim to develop compounds that maximize the effectiveness of the drug (the objective function) while minimizing toxicity (a constraint).

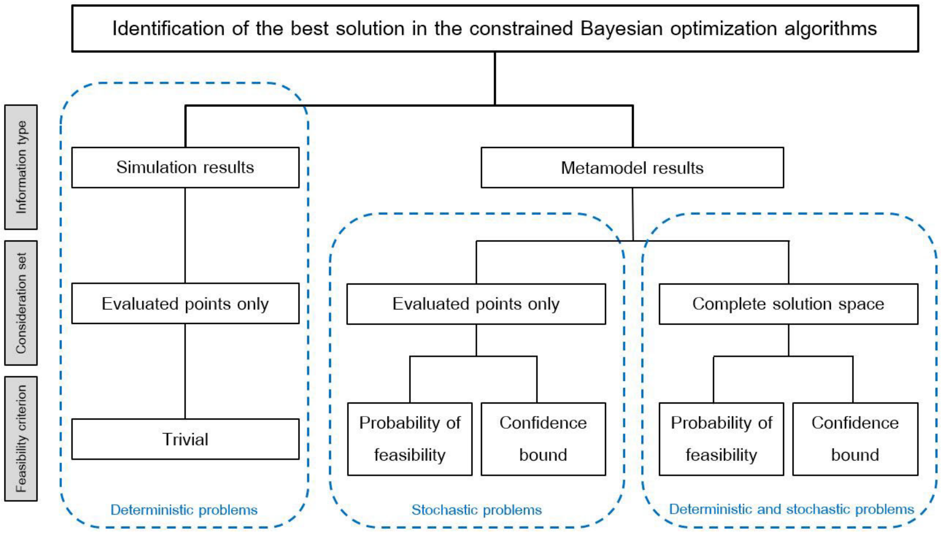

FIGURE 2. Identification methods in the CBO literature.

The decoupled setting allows these evaluations to be conducted independently. A compound's effectiveness can be tested first, and if it shows promise, toxicity tests can be run afterwards. This prevents unnecessary toxicity testing on ineffective compounds, saving time and resources, and thus improving efficiency in the drug development process.

Another example is the handling of mixed search spaces, which arise when the problem setting involves not only continuous but also discrete numerical and/or categorical variables. This commonly occurs in real-world settings: e.g., when designing an aircraft [78], continuous variables (such as wing span) are combined with categorical variables (such as engine type). Traditional GP-based BO methods are not naturally suited for this type of problems. The authors in [78] used a simple continuous relaxation approach, treating categorical variables as continuous during the optimization process.

A final example is the use of batch sampling in [73]. Batch sampling uses an acquisition function to select multiple points which are then evaluated in parallel. This may be more cost-effective and time-efficient than running sequential experiments, in particular in real-world problem settings that require substantial setup costs in order to perform evaluations (e.g., preparing equipment for physical experiments, initializing simulations, preparing an expensive testing environment, etc.). In industries like pharmaceutical testing or manufacturing, these setups are typically resource-intensive [99]. Batch sampling then reduces the number of setups required compared to single-experiment evaluations, leading to time and cost savings.

IV. CONCLUSION AND FURTHER RESEARCH DIRECTIONS

This paper surveys the current literature on single-objective constrained Bayesian optimization (CBO), focusing explicitly on three key algorithmic aspects: (*i*) the metamodels used, (*ii*) the acquisition function, and (*iii*) the identification procedure. The existing approaches are discussed, highlighting their advantages and drawbacks.

While fitting multiple surrogates is widely preferred by different authors, the potential benefits of parallel computation [100], [101] remain underexplored. Parallel computation in Bayesian optimization can alleviate the computational burden imposed by matrix inversion, particularly as the number of points increases. By distributing matrix inversion computations across multiple processing units (see e.g., [102]), parallelization reduces the time required for optimization.

So far, Gaussian processes have been the preferred metamodel in CBO algorithms. Yet, alternatives such as random forests [103] and neural networks [104] have been considered in the broader Bayesian optimization literature to enhance the performance of Bayesian algorithms. These alternatives may offer capabilities that are better aligned with problem-specific properties (e.g., mixed search space, scalability). It seems likely that these alternatives can take up a similar role also in CBO. In the broader metamodel-based optimization literature, *radial basis functions (RBF)* have also been used to solve constrained expensive black-box problems [105]. While Gaussian Processes are more advanced and provide better uncertainty quantification, certain features of radial basis functions, such as their ability to handle high-dimensional problems efficiently, as demonstrated in [106], can still be highly valuable. This opens the potential for a hybrid framework where the full capabilities of both GP and RBF are leveraged. Such a framework could benefit from the uncertainty modelling and interpretability of GP while utilizing the scalability and computational efficiency of RBF in high-dimensional settings. Recently, Tree-structured Parzen Estimators (TPE) have demonstrated strong performance in handling mixed search spaces, particularly in constrained optimization problems ([107]). The authors in [108] have already explored a hybrid framework combining GP and TPE, leveraging the strengths of both methods in solving (unconstrained)hyperparameter optimization problems.

Surprisingly, the large majority of research on CBO focuses on deterministic approaches (35 out of 48 surveyed papers in this review). Real-world problems are often stochastic though, and in constrained settings, the presence of noise introduces additional challenges for the optimization process. In the few papers that account for noise, the main effort has been spent on modifying the search and acquisition functions, yet the identification phase has been overlooked. Errors in the estimation of the PoF and/or the estimation of the goal value may prevent the algorithm from correctly identifying the true optimum in noisy settings, even when this optimal solution has been sampled during the search [45]. This highlights the need to develop methods that explicitly account for uncertainties in the identification phase.

Furthermore, the data efficiency of CBO might be enhanced further by implementing techniques that are currently gaining traction in the broader optimization field, such as multifidelity optimization or transfer learning. In multifidelity optimization [109], [110], both high- and low-fidelity evaluations can be made of the system at hand, with low-fidelity evaluations typically being less accurate but also cheaper to evaluate. Data can also be combined from disparate sources in view of improving the efficiency of the optimization; this is referred to as multi-information learning [111]. Transfer learning, on the other hand, is a machine learning technique that uses knowledge gained from solving one task to improve performance on another, yet similar task. It is especially useful when there is limited data for the target task but abundant data for a related source task. For instance, if a similar optimization problem has been tackled before, the metamodel can be initialized or informed using data from that previous task [112], reducing the number of function evaluations needed to reach an optimal solution. While transfer learning has been explored in unconstrained Bayesian optimization settings, extending it to CBO could be a valuable research direction to reduce the computational effort needed to handle complex constraints.

The papers reviewed in this work primarily focus on single-fidelity settings. Multifidelity optimization presents a promising direction for improving constrained Bayesian optimization (CBO). The authors of [109] applied multi-fidelity techniques to kriging models used in variable fidelity optimization, offering a first attempt at reducing computational costs by integrating low- and high-fidelity models. A more recent paper [110], compares a mono-fidelity Bayesian optimization method with its multi-fidelity counterpart on a drone design optimization problem. While multi-fidelity typically refers to a hierarchical relationship among models, where higher fidelity is always more accurate but also more expensive to evaluate, [111] consider a multi-information setting where various sources of information may not have a clear fidelity hierarchy. In this setting, the focus is on fusing data from disparate sources, irrespective of whether there is a fidelity relationship between them, and efficiently using this combined information for optimization. While these approaches showed potential, further advancements could improve the handling of complex constraints and the balance between model accuracy and computational efficiency. Future research could refine these techniques, particularly in optimizing the use of fidelity models to enhance convergence speed and data efficiency in CBO.

As another potential research direction, the application of transfer learning [113] in constrained Bayesian optimization (CBO) offers a promising avenue for improving data efficiency by leveraging prior knowledge from related optimization tasks. Transfer learning is a machine learning technique where knowledge gained from solving one task is transferred and reused to improve performance on a different but similar task. It is especially useful when there is limited data for the target task but abundant data for a related source task. For instance, if a similar optimization problem has been tackled before, the metamodel can be initialized or informed using data from that previous task [112], reducing the number of function evaluations needed to reach an optimal solution. While transfer learning has been explored in unconstrained Bayesian optimization settings, extending it to CBO could be a valuable research direction to reduce the computational effort needed to handle complex constraints.

Finally, in the current CBO literature, the only focus is on finding a solution that optimizes the expected performance of the objective function of interest, given the constraint requirements. Optimizing the expected performance of the system is typically relevant for risk-neutral decision-makers. In practice, decision-makers tend to be risk-averse and are thus primarily willing to give in on the expected result in return for a decrease in the risk/uncertainty. This research field is known as robust Bayesian optimization [114], [115]. In the unconstrained BO literature, two recent papers have focused on optimizing a risk measure, i.e., the value-at-risk [115] or the conditional value-at-risk [114]. Developing similar methods for constrained settings thus remains an interesting and highly relevant research direction.

REFERENCES

[1] R. Garnett, *Bayesian Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2023.

[2] X. Wang, Y. Jin, S. Schmitt, and M. Olhofer, "Recent advances in Bayesian optimization," *ACM Comput. Surv.*, vol. 55, no. 13s, pp. 1–36, Dec. 2023.

[3] P. I. Frazier, "A tutorial on Bayesian optimization," 2018, *arXiv:1807.02811*.

[4] A. Candelieri, "A gentle introduction to Bayesian optimization," in *Proc. Winter Simulation Conf. (WSC)*, Dec. 2021, pp. 1–16.

[5] A. Morales-Hernández, S. R. Gonzalez, I. Van Nieuwenhuyse, I. Couckuyt, J. Jordens, M. Witters, and B. Van Doninck, "Bayesian multi-objective optimization of process design parameters in constrained settings with noise: An engineering design application," *Eng. Comput.*, vol. 40, no. 4, pp. 2497–2511, Aug. 2024.

[6] N. Loka, I. Couckuyt, F. Garbuglia, D. Spina, I. Van Nieuwenhuyse, and T. Dhaene, "Bi-objective Bayesian optimization of engineering problems with cheap and expensive cost functions," *Eng. Comput.*, vol. 39, no. 3, pp. 1923–1933, Jun. 2023.

[7] N. Loka, M. Ibrahim, I. Couckuyt, I. Van Nieuwenhuyse, and T. Dhaene, "Cheap-expensive multi-objective Bayesian optimization for permanent magnet synchronous motor design," *Eng. Comput.*, vol. 40, no. 4, pp. 2143–2159, Aug. 2024.

[8] Y. Morita, S. Rezaeiravesh, N. Tabatabaei, R. Vinuesa, K. Fukagata, and P. Schlatter, "Applying Bayesian optimization with Gaussian process regression to computational fluid dynamics problems," *J. Comput. Phys.*, vol. 449, Jan. 2022, Art. no. 110788.

[9] S. Park, J. Na, M. Kim, and J. M. Lee, "Multi-objective Bayesian optimization of chemical reactor design using computational fluid dynamics," *Comput. Chem. Eng.*, vol. 119, pp. 25–37, Nov. 2018.

[10] B. J. Shields, J. Stevens, J. Li, M. Parasram, F. Damani, J. I. M. Alvarado, J. M. Janey, R. P. Adams, and A. G. Doyle, "Bayesian reaction optimization as a tool for chemical synthesis," *Nature*, vol. 590, no. 7844, pp. 89–96, Feb. 2021.

[11] G. E. P. Box and G. C. Tiao, *Bayesian Inference in Statistical Analysis*. Hoboken, NJ, USA: Wiley, 2011.

[12] R. Lam, M. Poloczek, P. Frazier, and K. E. Willcox, "Advances in Bayesian optimization with applications in aerospace engineering," in *Proc. AIAA Non-Deterministic Approaches Conf.*, Jan. 2018, p. 1656.

[13] H. M. Torun, M. Swaminathan, A. K. Davis, and M. L. F. Bellaredj, "A global Bayesian optimization algorithm and its application to integrated system design," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 26, no. 4, pp. 792–802, Apr. 2018.

[14] F. Archetti and A. Candelieri, *Bayesian Optimization and Data Science*, vol. 849. Cham, Switzerland: Springer, 2019.

[15] B. Letham, B. Karrer, G. Ottoni, and E. Bakshy, "Constrained Bayesian optimization with noisy experiments," *Bayesian Anal.*, vol. 14, no. 2, pp. 495–519, Jun. 2019.

[16] J. M. Hernández-Lobato, M. A. Gelbart, R. P. Adams, M. W. Hoffman, and Z. Ghahramani, "A general framework for constrained Bayesian optimization using information-based search," *J. Mach. Learn. Res.*, vol. 17, no. 1, pp. 5549–5601, Jan. 2016.

[17] R. B. Gramacy and H. K. H. Lee, "Optimization under unknown constraints," in *Proc. 9th Bayesian Statist. Int. Meeting*, Jan. 2010, pp. 229–256.

[18] F. Berkenkamp, A. Krause, and A. P. Schoellig, "Bayesian optimization with safety constraints: Safe and automatic parameter tuning in robotics," *Mach. Learn.*, vol. 112, no. 10, pp. 3713–3747, Oct. 2023.

[19] Y. Sui, V. Zhuang, J. W. Burdick, and Y. Yue, "Stagewise safe Bayesian optimization with Gaussian processes," in *Proc. Int. Conf. Mach. Learn.*, Jan. 2018, pp. 4781–4789.

[20] K. M. Atkinson, A. C. Koenka, C. E. Sanchez, H. Moshontz, and H. Cooper, "Reporting standards for literature searches and report inclusion criteria: Making research syntheses more transparent and easy to replicate," *Res. Synth. Methods*, vol. 6, no. 1, pp. 87–95, Mar. 2015.

[21] 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, Jan. 2016.

[22] P. I. Frazier, "Bayesian optimization," in *Recent Advances in Optimization and Modeling of Contemporary Problems*. USA: INFORMS, 2018, pp. 255–278, doi: 10.1287/educ.2018.0188.

[23] S. Greenhill, S. Rana, S. Gupta, P. Vellanki, and S. Venkatesh, "Bayesian optimization for adaptive experimental design: A review," *IEEE Access*, vol. 8, pp. 13937–13948, 2020.

[24] R. G. Regis, "A survey of surrogate approaches for expensive constrained black-box optimization," in *Optimization of Complex Systems: Theory, Models, Algorithms and Applications*. Cham, Switzerland: Springer, 2020, pp. 37–47.

[25] S. Le Digabel and S. M. Wild, "A taxonomy of constraints in black-box simulation-based optimization," *Optim. Eng.*, vol. 25, no. 2, pp. 1125–1143, Jun. 2024.

[26] B. Aczel, R. Hoekstra, A. Gelman, E.-J. Wagenmakers, I. G. Klugkist, J. N. Rouder, J. Vandekerckhove, M. D. Lee, R. D. Morey, W. Vanpaemel, Z. Dienes, and D. van Ravenzwaaij, "Discussion points for Bayesian inference," *Nature Hum. Behav.*, vol. 4, no. 6, pp. 561–563, Jan. 2020.

[27] D. R. Jones, M. Schonlau, and W. J. Welch, "Efficient global optimization of expensive black-box functions," *J. Global Optim.*, vol. 13, no. 4, pp. 455–492, Dec. 1998.

[28] V. Picheny, T. Wagner, and D. Ginsbourger, "A benchmark of kriging-based infill criteria for noisy optimization," *Struct. Multidisciplinary Optim.*, vol. 48, no. 3, pp. 607–626, Sep. 2013.

[29] S. Tautvaišas and J. Žilinskas, "Scalable Bayesian optimization with generalized product of experts," *J. Global Optim.*, vol. 88, no. 3, pp. 777–802, Mar. 2024.

[30] J. Qing, I. Couckuyt, and T. Dhaene, "A robust multi-objective Bayesian optimization framework considering input uncertainty," *J. Global Optim.*, vol. 86, no. 3, pp. 693–711, Jul. 2023.

[31] S. Tautvaišas and J. Žilinskas, "Heteroscedastic Bayesian optimization using generalized product of experts," *J. Global Optim.*, pp. 1–21, Nov. 2023, doi: 10.1007/s10898-023-01333-5.

[32] M. A. Gelbart, J. Snoek, and R. P. Adams, "Bayesian optimization with unknown constraints," 2014, *arXiv:1403.5607*.

[33] 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.

[34] B. Ankenman, B. L. Nelson, and J. Staum, "Stochastic Kriging for simulation metamodeling," *Oper. Res.*, vol. 58, no. 2, pp. 371–382, Apr. 2010, doi: 10.1287/opre.1090.0754.

[35] S. Rojas-Gonzalez and I. Van Nieuwenhuyse, "A survey on kriging-based infill algorithms for multiobjective simulation optimization," *Comput. Operations Res.*, vol. 116, Apr. 2020, Art. no. 104869.

[36] H. J. Kushner, "A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise," *J. Basic Eng.*, vol. 86, no. 1, pp. 97–106, Mar. 1964, doi: 10.1115/1.3653121.

[37] D. Zhan and H. Xing, "Expected improvement for expensive optimization: A review," *J. Global Optim.*, vol. 78, no. 3, pp. 507–544, Nov. 2020.

[38] R. Priem, N. Bartoli, Y. Diouane, and A. Sgueglia, "Upper trust bound feasibility criterion for mixed constrained Bayesian optimization with application to aircraft design," *Aerosp. Sci. Technol.*, vol. 105, Oct. 2020, Art. no. 105980.

[39] Q. Nguyen, *Bayesian Optimization in Action*. Shelter Island, NY, USA: MANNING, 2023.

[40] T. Pourmohamad and H. K. H. Lee, "Bayesian optimization via barrier functions," *J. Comput. Graph. Statist.*, vol. 31, no. 1, pp. 74–83, Jan. 2022.

[41] C. Doerr, "Complexity theory for discrete black-box optimization heuristics," in *Theory of Evolutionary Computation*. Cham, Switzerland: Springer, 2020, pp. 133–212.

[42] J. E. Fieldsend and R. M. Everson, "The rolling tide evolutionary algorithm: A multiobjective optimizer for noisy optimization problems," *IEEE Trans. Evol. Comput.*, vol. 19, no. 1, pp. 103–117, Feb. 2015.

[43] J. Yin, S. H. Ng, and K. M. Ng, "Kriging metamodel with modified nugget-effect: The heteroscedastic variance case," *Comput. Ind. Eng.*, vol. 61, no. 3, pp. 760–777, Oct. 2011.

[44] X. Chen, X. He, C. Xiong, Z. Zhu, and L. Zhang, "A Bayesian stochastic Kriging optimization model dealing with heteroscedastic simulation noise for freeway traffic management," *Transp. Sci.*, vol. 53, no. 2, pp. 545–565, Mar. 2019.

[45] S. Amini and I. Van Nieuwenhuyse, "A Bayesian optimization algorithm for constrained simulation optimization problems with heteroscedastic noise," in *Proc. Int. Conf. Learn. Intell. Optim.* Cham, Switzerland: Springer, Jan. 2023, pp. 78–91.

[46] Z. Wang and M. Ierapetritou, "Constrained optimization of black-box stochastic systems using a novel feasibility enhanced Kriging-based method," *Comput. Chem. Eng.*, vol. 118, pp. 210–223, Oct. 2018.

[47] R. B. Gramacy, G. A. Gray, S. Le Digabel, H. K. H. Lee, P. Ranjan, G. Wells, and S. M. Wild, "Modeling an augmented Lagrangian for blackbox constrained optimization," *Technometrics*, vol. 58, no. 1, pp. 1–11, Jan. 2016.

[48] T. Pourmohamad and H. K. H. Lee, "The statistical filter approach to constrained optimization," *Technometrics*, vol. 62, no. 3, pp. 303–312, Jul. 2020.

[49] T. Pourmohamad and H. K. H. Lee, "Multivariate stochastic process models for correlated responses of mixed type," *Bayesian Anal.*, vol. 11, no. 3, pp. 797–820, Sep. 2016.

[50] J. R. Gardner, M. J. Kusner, Zhixiang, K. Q. Weinberger, and J. P. Cunningham, "Bayesian optimization with inequality constraints," in *Proc. ICML*, Jun. 2014, pp. 937–945.

[51] J. Ungredda and J. Branke, "Bayesian optimisation for constrained problems," *ACM Trans. Model. Comput. Simul.*, vol. 34, no. 2, pp. 1–26, Apr. 2024.

[52] S. Zhang, R. M. Lee, B. Shafei, D. Walz, and R. Misener, "Dependence in constrained Bayesian optimization: When do we need it and how does it help?" *Optim. Lett.*, vol. 18, no. 6, pp. 1–17, 2023.

[53] R. R. Carpio, R. C. Giordano, and A. R. Secchi, ''Enhanced surrogate assisted framework for constrained global optimization of expensive black-box functions,'' *Comput. Chem. Eng.*, vol. 118, pp. 91–102, Oct. 2018.
[54] V. Picheny, D. Ginsbourger, O. Roustant, R. T. Haftka, and N.-H. Kim, ''Adaptive designs of experiments for accurate approximation of a target region,'' *J. Mech. Design*, vol. 132, no. 7, Jul. 2010.
[55] K. Miettinen, *Nonlinear Multiobjective Optimization*, vol. 12. Cham, Switzerland: Springer, 2012.
[56] M. J. Sasena, P. Papalambros, and P. Goovaerts, ''Exploration of metamodeling sampling criteria for constrained global optimization,'' *Eng. Optim.*, vol. 34, no. 3, pp. 263–278, Jan. 2002.
[57] H. Lee, R. Gramacy, C. Linkletter, and G. Gray, ''Optimization subject to hidden constraints via statistical emulation,'' *Pacific J. Optim.*, vol. 7, no. 3, pp. 467–478, 2011.
[58] J. M. Parr, A. J. Keane, A. I. J. Forrester, and C. M. E. Holden, ''Infill sampling criteria for surrogate-based optimization with constraint handling,'' *Eng. Optim.*, vol. 44, no. 10, pp. 1147–1166, Oct. 2012.
[59] V. Picheny, ''A stepwise uncertainty reduction approach to constrained global optimization,'' in *Proc. Artif. Intell. Statist.*, Apr. 2014, pp. 787–795.
[60] D. V. Lindberg and H. K. H. Lee, ''Optimization under constraints by applying an asymmetric entropy measure,'' *J. Comput. Graph. Statist.*, vol. 24, no. 2, pp. 379–393, Apr. 2015.
[61] C. Durantin, J. Marzat, and M. Balesdent, ''Analysis of multi-objective kriging-based methods for constrained global optimization,'' *Comput. Optim. Appl.*, vol. 63, no. 3, pp. 903–926, Apr. 2016.
[62] V. Picheny, R. B. Gramacy, S. M. Wild, and S. L. Digabel, ''Bayesian optimization under mixed constraints with a slack-variable augmented Lagrangian,'' in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 29, Jan. 2016.
[63] S. Bagheri, W. Konen, R. Allmendinger, J. Branke, K. Deb, J. Fieldsend, D. Quagliarella, and K. Sindhya, ''Constraint handling in efficient global optimization,'' in *Proc. Genetic Evol. Comput. Conf.*, Jul. 2017, pp. 673–680.
[64] R. Lam and K. Willcox, ''Lookahead Bayesian optimization with inequality constraints,'' in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, Jan. 2017, pp. 1890–1900.
[65] Y. Li, Y. Wu, J. Zhao, and L. Chen, ''A Kriging-based constrained global optimization algorithm for expensive black-box functions with infeasible initial points,'' *J. Global Optim.*, vol. 67, nos. 1–2, pp. 343–366, Jan. 2017.
[66] I.-B. Chung, D. Park, and D.-H. Choi, ''Surrogate-based global optimization using an adaptive switching infill sampling criterion for expensive black-box functions,'' *Struct. Multidisciplinary Optim.*, vol. 57, no. 4, pp. 1443–1459, Apr. 2018.
[67] H. Dong, B. Song, Z. Dong, and P. Wang, ''SCGOSR: Surrogate-based constrained global optimization using space reduction,'' *Appl. Soft Comput.*, vol. 65, pp. 462–477, Apr. 2018.
[68] B. Yuan, L. Liu, T. Long, and R. Shi, ''Efficient global optimization strategy considering expensive constraints,'' in *Proc. 12th World Congr. Struct. Multidisciplinary Optim. (WCSMO)*. Cham, Switzerland: Springer, Dec. 2017, pp. 133–142.
[69] S. Ariafar, J. Coll-Font, D. H. Brooks, and J. Dy, ''ADMMBO: Bayesian optimization with unknown constraints using ADMM,'' *J. Mach. Learn. Res.*, vol. 20, pp. 1–26, Jan. 2019.
[70] N. Bartoli, T. Lefebvre, S. Dubreuil, R. Olivanti, R. Priem, N. Bons, J. R. R. A. Martins, and J. Morlier, ''Adaptive modeling strategy for constrained global optimization with application to aerodynamic wing design,'' *Aerosp. Sci. Technol.*, vol. 90, pp. 85–102, Jul. 2019.
[71] R. Jiao, S. Zeng, C. Li, Y. Jiang, and Y. Jin, ''A complete expected improvement criterion for Gaussian process assisted highly constrained expensive optimization,'' *Inf. Sci.*, vol. 471, pp. 80–96, Jan. 2019.
[72] R. Shi, L. Liu, T. Long, Y. Wu, and Y. Tang, ''Filter-based adaptive Kriging method for black-box optimization problems with expensive objective and constraints,'' *Comput. Methods Appl. Mech. Eng.*, vol. 347, pp. 782–805, Apr. 2019.
[73] A. Tran, J. Sun, J. M. Furlan, K. V. Pagalthivarthi, R. J. Visintainer, and Y. Wang, ''PBO-2GP-3B: A batch parallel known/unknown constrained Bayesian optimization with feasibility classification and its applications in computational fluid dynamics,'' *Comput. Methods Appl. Mech. Eng.*, vol. 347, pp. 827–852, Apr. 2019.
[74] H. Akbari and A. Kazerooni, ''KASRA: A Kriging-based adaptive space reduction algorithm for global optimization of computationally expensive black-box constrained problems,'' *Appl. Soft Comput.*, vol. 90, May 2020, Art. no. 106154.
[75] A. Parnianifard, R. Chancharoen, G. Phanomchoeng, and L. Wuttisittikulkij, ''A new approach for low-dimensional constrained engineering design optimization using design and analysis of simulation experiments,'' *Int. J. Comput. Intell. Syst.*, vol. 13, no. 1, p. 1663, 2020.
[76] T. Tao, G. Zhao, and S. Ren, ''An efficient kriging-based constrained optimization algorithm by global and local sampling in feasible region,'' *J. Mech. Design*, vol. 142, no. 5, May 2020.
[77] P. Jiang, Y. Cheng, J. Yi, and J. Liu, ''An efficient constrained global optimization algorithm with a clustering-assisted multiobjective infill criterion using Gaussian process regression for expensive problems,'' *Inf. Sci.*, vol. 569, pp. 728–745, Aug. 2021.
[78] P. Saves, N. Bartoli, Y. Diouane, T. Lefèbvre, J. Morlier, C. David, E. N. Van, and S. Defoort, ''Constrained Bayesian optimization over mixed categorical variables, with application to aircraft design,'' in *Proc. AeroBest*, Jul. 2021.
[79] D. Siefman, C. Percher, J. Norris, A. Kersting, and D. Heinrichs, ''Constrained Bayesian optimization of criticality experiments,'' *Ann. Nucl. Energy*, vol. 151, Feb. 2021, Art. no. 107894.
[80] Y. Zhang, X. Zhang, and P. I. Frazier, ''Constrained two-step look-ahead Bayesian optimization,'' in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 34, 2021, pp. 12563–12575.
[81] S. Zhang, F. Yang, C. Yan, D. Zhou, and X. Zeng, ''An efficient batch-constrained Bayesian optimization approach for analog circuit synthesis via multiobjective acquisition ensemble,'' *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 41, no. 1, pp. 1–14, Jan. 2022.
[82] Y. Zeng, Y. Cheng, and J. Liu, ''An efficient global optimization algorithm for expensive constrained black-box problems by reducing candidate infilling region,'' *Inf. Sci.*, vol. 609, pp. 1641–1669, Sep. 2022.
[83] Q. P. Nguyen, W. T. R. Chew, Le Song, B. K. H. Low, and P. Jaillet, ''Optimistic Bayesian optimization with unknown constraints,'' in *Proc. 12th Int. Conf. Learn. Represent.*, 2023.
[84] W. Xu, Y. Jiang, B. Svetozarevic, and C. N. Jones, ''Constrained efficient global optimization of expensive black-box functions,'' in *Proc. Int. Conf. Mach. Learn.*, Jan. 2022, pp. 38485–38498.
[85] C. Bian, Q. Liu, X. Zhang, B. Yan, X. Wang, S. Zuo, and H. Liu, ''An efficient mixed constrained Bayesian optimization for handling known and unknown constraints,'' *Adv. Eng. Informat.*, vol. 62, Oct. 2024, Art. no. 102704.
[86] J. Pelamatti, R. Le Riche, C. Helbert, and C. Blanchet-Scalliet, ''Coupling and selecting constraints in Bayesian optimization under uncertainties,'' *Optim. Eng.*, vol. 25, no. 1, pp. 373–412, Mar. 2024.
[87] Q. Ragueneau, L. Laurent, A. Legay, T. Larroque, and R. Crambuer, ''A constrained Bayesian optimization framework for structural vibrations with local nonlinearities,'' *Struct. Multidisciplinary Optim.*, vol. 67, no. 4, p. 47, Apr. 2024.
[88] J. Song, Y. Cui, P. Wei, M. A. Valdebenito, and W. Zhang, ''Constrained Bayesian optimization algorithms for estimating design points in structural reliability analysis,'' *Rel. Eng. Syst. Saf.*, vol. 241, Jan. 2024, Art. no. 109613.
[89] A. Tfaily, Y. Diouane, N. Bartoli, and M. Kokkolaras, ''Bayesian optimization with hidden constraints for aircraft design,'' *Struct. Multidisciplinary Optim.*, vol. 67, no. 7, p. 123, Jul. 2024.
[90] C. A. C. Coello, ''Constraint-handling techniques used with evolutionary algorithms,'' in *Proc. Genetic Evol. Comput. Conf. Companion*, Jul. 2019, pp. 1310–1333.
[91] K. Deb, K. Sindhya, and J. Hakanen, *Multi-Objective Optimization*. Boca Raton, FL, USA: CRC Press, 2016, pp. 161–200.
[92] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, ''A fast and elitist multiobjective genetic algorithm: NSGA-II,'' *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, Apr. 2002.
[93] S. Amini, ''Constrained Bayesian optimization with barrier functions; impact of heteroscedastic noise on the search and identification phases,'' Ph.D. thesis, Data Sci. Inst., Hasselt Univ., Hasselt, Belgium, 2024.
[94] L. Mihaela Paun, A. Fensterseifer Schmidt, S. McGinty, and D. Husmeier, ''Constrained Bayesian optimization with a cardiovascular application,'' *Proc. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 480, no. 2295, Aug. 2024, Art. no. 20230371.
[95] M. Binois and N. Wycoff, ''A survey on high-dimensional Gaussian process modeling with application to Bayesian optimization,'' *ACM Trans. Evol. Learn. Optim.*, vol. 2, no. 2, pp. 1–26, Jun. 2022.
[96] D. Eriksson and M. Jankowiak, ''High-dimensional Bayesian optimization with sparse axis-aligned subspaces,'' in *Proc. Uncertainty Artif. Intell.*, Jan. 2021, pp. 493–503.

[97] R. Moriconi, M. P. Deisenroth, and K. S. Sesh Kumar, "High-dimensional Bayesian optimization using low-dimensional feature spaces," *Mach. Learn.*, vol. 109, nos. 9–10, pp. 1925–1943, Sep. 2020.
[98] L. Colliandre and C. Müller, "Bayesian optimization in drug discovery," in *High Performance Computing for Drug Discovery and Biomedicine*. Cham, Switzerland: Springer, 2023, pp. 101–136.
[99] G. Schneider, "Automating drug discovery," *Nature Rev. Drug Discovery*, vol. 17, no. 2, pp. 97–113, Feb. 2018.
[100] J. Wu and P. Frazier, "The parallel knowledge gradient method for batch Bayesian optimization," in *Proc. Adv. Neural Inf. Process. Syst.*, 2016.
[101] L. D. González and V. M. Zavala, "New paradigms for exploiting parallel experiments in Bayesian optimization," *Comput. Chem. Eng.*, vol. 170, Feb. 2023, Art. no. 108110.
[102] D. Yu, S. He, Y. Huang, G. Yu, and L. Yang, "A fast parallel matrix inversion algorithm based on heterogeneous multicore architectures," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, Dec. 2015, pp. 903–907.
[103] Z. Wang, F. Hutter, M. Zoghi, D. Matheson, and N. De Feitas, "Bayesian optimization in a billion dimensions via random embeddings," *J. Artif. Intell. Res.*, vol. 55, pp. 361–387, Feb. 2016.
[104] A. Klein, S. Falkner, S. Bartels, P. Hennig, and F. Hutter, "Fast Bayesian optimization of machine learning hyperparameters on large datasets," in *Proc. Artif. Intell. Statist.*, Jan. 2016, pp. 528–536.
[105] R. G. Regis, "Stochastic radial basis function algorithms for large-scale optimization involving expensive black-box objective and constraint functions," *Comput. Oper. Res.*, vol. 38, no. 5, pp. 837–853, May 2011.
[106] R. G. Regis, "Constrained optimization by radial basis function interpolation for high-dimensional expensive black-box problems with infeasible initial points," *Eng. Optim.*, vol. 46, no. 2, pp. 218–243, Feb. 2014.
[107] S. Watanabe and F. Hutter, "C-TPE: Tree-structured Parzen estimator with inequality constraints for expensive hyperparameter optimization," in *Proc. 32nd Int. Joint Conf. Artif. Intell.*, Aug. 2023, pp. 4371–4379.
[108] A. M. Hernandez, I. Van Nieuwenhuyse, and S. R. Gonzalez, "Tree-structured Parzen estimators with uncertainty for hyperparameter optimization of machine learning algorithms," in *Proc. Winter Simulation Conf.*, 2022.
[109] S. E. Gano, J. E. Renaud, J. D. Martin, and T. W. Simpson, "Update strategies for Kriging models used in variable fidelity optimization," *Struct. Multidisciplinary Optim.*, vol. 32, no. 4, pp. 287–298, Oct. 2006.
[110] R. Charayron, T. Lefèbvre, N. Bartoli, and J. Morlier, "Multi-fidelity constrained Bayesian optimization, application to drone design," in *Proc. Conf. Artif. Intell. Defense CAID*, Nov. 2021.
[111] S. F. Ghoreishi and D. Allaire, "Multi-information source constrained Bayesian optimization," *Structural Multidisciplinary Optim.*, vol. 59, no. 3, pp. 977–991, Mar. 2019.
[112] P. Tighineanu, K. Skubch, P. Baireuther, A. Reiss, F. Berkenkamp, and J. Vinogradska, "Transfer learning with Gaussian processes for Bayesian optimization," in *Proc. Int. Conf. Artif. Intell. Statist.*, Jan. 2021, pp. 6152–6181.
[113] T. Bai, Y. Li, Y. Shen, X. Zhang, W. Zhang, and B. Cui, "Transfer learning for Bayesian optimization: A survey," 2023, *arXiv:2302.05927*.
[114] Q. P. Nguyen, Z. Dai, B. K. H. Low, and P. Jaillet, "Optimizing conditional value-at-risk of black-box functions," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 34, Dec. 2021, pp. 4170–4180.
[115] Q. P. Nguyen, Z. Dai, B. K. H. Low, and P. Jaillet, "Value-at-risk optimization with Gaussian processes," in *Proc. Int. Conf. Mach. Learn.*, Jan. 2021, pp. 8063–8072.

SASAN AMINI received the bachelor's and master's degrees in industrial engineering, and the Ph.D. degree in business economics from Hasselt University, Belgium. He is currently a Postdoctoral Researcher with the Computational Mathematics Research Group, Hasselt University. His research interests include constraint handling methods in black-box optimization problems, with a particular emphasis on stochastic settings and problems involving expensive-to-evaluate functions.

INNEKE VANNIEUWENHUYSE is currently a Professor with Hasselt University and KU Leuven, Belgium. Her research interests include (multi-objective) optimization of stochastic systems, especially in settings that are expensive to evaluate. To that end, she combines operations research techniques with machine learning approaches.

ALEJANDRO MORALES-HERNÁNDEZ received the Ph.D. degree in business economics from Hasselt University, and the Ph.D. degree in computer science engineering from Ghent University. He is currently a Postdoctoral Research Fellow with the Machine Learning Group, Université Libre de Bruxelles, Belgium. Since 2019, his research work has been centered on developing stochastic optimization algorithms assisted by machine learning techniques for single and multi-objective decision problems, with a focus on noise-handling strategies using Bayesian learning methods. Recently he has focused on time series analysis and forecasting in the context of green energy and vehicular traffic.