Codimension-Two
Free Boundary Problems

Keith Gillow

St Catherine's College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 1998

This thesis is dedicated to
someone
for some special reason

Acknowledgements

plenty of waffle, plenty of waffle, plenty of waffle, plenty of waffle, plenty of waffle, plenty of waffle, plenty of waffle, plenty of waffle.

Abstract

plenty of waffle, plenty of waffle, plenty of waffle, plenty of waffle, plenty of waffle, plenty of waffle, plenty of waffle, plenty of waffle.

Contents

List of Figures

Chapter 1

Introduction

Bayesian optimisation is considered to be the specific case of model-based optimisation methods which is structured based on the Bayesian formulations. Like the most cases of Bayesian approaches, there is a prior distribution over a function, a likelihood function, and a posterior distribution over the unknown functions given data [16]. Bayesian optimisation aims to solve a difficult though concise problem:

$$\min_{x \in \mathcal{X}} f(x) \tag{1.1}$$

However $f(x)$ is considered to have following features [16]:

1. **Expensive to Evaluate:** Evaluation of $f(x)$ is highly expensive. Such as most of engineering problems like making alloys [39], desigining a pharmaceutical drug, optimising machine learning algorithms [35] and many other engineering problems.

2. **Black-box:** The exact formulation of $f(x)$ function or the derivatives of $f(x)$ is not available. Like the most of the engineering problems we have already mentioned.

3. **Noisy Evaluations:** The value of $f(x)$ for a specific x usually corrupted with a noise. So for the same values of x, $f(x)$ would result in different values.

4. **Global and non-convex:** Bayesian optimization of $f(x)$ is a global optimisation problem for a non-convex function in the $\mathcal{X}$ domain.

While traditional numerical methods have proved ineffective for solving some optimisation problems, Bayesian optimisation has proved to be effective in variety of optimisation problems dealing with black-box objective functions expensive to evaluate.

1

There have been number of studies on the use of Bayesian optimisation on hyper-parameter tuning in machine learning and big data [23], expensive multi-objective optimisation for Robotics [37], and experimentation optimisation in product design such as short polymer fiber materials [25].

Though practical problems are often involved in several non-commensurable objectives. In other words, the real-world problems consist of multiple, conflicting black-box objectives. In Multi-Objective Optimisation (MOO) potential solutions are assessed by their performance in more than one objective [7]. In MOO, based on definition of Pareto optimality, we wish to return a Pareto front that represents the best trade-off possible considering all criteria [5]. More generally, MOO includes M objective functions $f_1, f_2, \ldots, f_M$ which are usually modelled by Gaussian Process (GP) [32]. Formally:

$$\begin{aligned}
\underset{\mathbf{x}}{\text{minimize}} \quad & \mathbf{y} = \mathbf{Z}(\mathbf{x}) = \{f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_M(\mathbf{x})\} \\
\text{where:} \quad & \mathbf{x} \in \mathcal{X} \subseteq \mathbb{R}^d \\
& \mathbf{y} \in \mathcal{Y} \subseteq \mathbb{R}^M
\end{aligned}$$

Generally two categoreis of solutions are available for solving multi-objective problems. In such scenarios it is customary to seek for a set of Pareto optimal outcomes often called the Pareto front [7]. The first category uses scalarization for transforming a multi-objective problem into single-objective problem. The single-objective problem is then tackled using different evolutionary algorithms such as Genetics Algorithm. So in this case, such solutions are not suited for the highly expensive and black-box $f(x)$ function. But there are studies trying to solve the transformed single-objective problem by Bayesian optimisation or entropy search methods. In this case, researchers try to eliminate the highly expensive costs of evolutionary algorithms. In such studies, there are differnt methods of scalarization in order to solve the multi-objective problems. Tchebycheff Method, k^{th}-Objective Weighted-Constraint, and Pascoletti and Serafini Scalarization are the most three popular approaches for scalarization among researchers [30].

The next category of solutions are focused on the multi-objective structure of the problem. These methods do not transform the multi-objective problems into single-objective space. Instead, these methods tend to use more complex acquisition functions based on Expected Hypervolume Improvement or Entropy Search in order to handle the multidimensionality of the problem. So they provide a set of solutions by optimizing all M objective simultaneously.

1.1   Motivation

Large diversity of optimisation problems take the form described above. Examples include increasing the Pouring Temperature of an alloy to withstand enormously high temperature of molted metals while overall alloy hardness must not increase drastically; or tuning training parameters of an SVM model to maximize the accuracy and minimizing the consumption of resources. In addition to expensive evaluation of objective functions, many optimisation problems deal with similarly expensive black-box constraints. Unknown constraints are part of many black-box multi-objective optimisation problems. For example, when tuning SVM hyperparameters we may want to optimise performance subject to a limit on the number of support vectors (and hence the complexity of evaluating the trained classifier) if the trained machine is to be implemented on limited hardware (such as accessible memory). The goal of optimisation in such cases is to minimize the number of black-box function evaluations to find the global optimum of the function with respect to constraints. Bayesian optimisation has recently been used successfully in this area [1]. Another use of multi-objective with constraints problem arises when there is a priori such as Ranking of the objectives. In many real-world problems, there are more important objectives we would like to have a slight advantage over other objectives. For example, in autonomous flying airplanes, the nature of the problem demands a huge advantage on safety measures of the airplane than the fuel consumption; or in Robotics, one of the current problems researches are working on, is about overheating of motors or power consumption of them. While other objectives such as the accuracy of movements with some constraints are also play a role in this problem. One may like to focus on the safety of motors or the accuracy of the movements. So ranking of the objectives as a priori information could be really useful in many real-world problems.

1.2   Summary of contributions

The main contribution of our first proposed method is to characterrize a general formulation for Multi-objective Bayesian optimisation with unknown constraints based on Hypervolume calculation. The other realted contributions are:

- Formulation of the expected hypervolume improvement with constraints based on the simple but effective expected improvement acquisition function.

- Evaluation of the proposed algorithm based on feasible dominated region on all related benchmark test functions for the first time. We also estimated the volume of the feasible region of the test functions for more accurate evaluation.

- Discussion of the issues involved in the method in terms of the efficiency and size of the problem.

The main contribution of multi-objective Bayesian optimisation with constraints and objective rankings is incorporating the rankings as a feature to multi-objective Bayesian optimisation. The contributions of the proposed model are summarized in the following three aspects:

- We formulated a single-objective constrained Bayesian optimisation problem for mapping a multi-objective Bayesian optimisation with constraints and objective rankings into feasible space.

- For evaluation of the proposed algorithm, we will introduce new measurements which should contain both diversity/density and accuracy of the obtained Pareto set.

- It is the first time that multi-objective Bayesian optimisation has been incorporated in ranking or planning problems.

Chapter 2

Literature Review

Bayesian optimisation is a well-known tool for solving a variety of optimisation problems. While traditional numerical methods have proved ineffective for solving some optimisation problems, Bayesian optimisation has proved in variety of optimisation problems dealing with blackbox objective functions expensive to evaluate [15]. Bayesian optimization is impacting a wide range of areas, including Robotics [26, 28], enviromental monitoring [27], interactive user interface [2], information extraction [40], combinatorial optimisation [41, 21], reinforcement learning [3], sensor networks [13, 36], and automated machine learning algorithms [38, 35, 18, 12].

Fundamentally, Bayesian optimization is a sequential model-based approach for solving optimisation problems. Bayesian optimisation framework has two main stages. The first one is a probabilistic surrogate model, consist of a prior distribution which encodes our beliefs about the nature of the expensive black-box function [34]. The second stage is constructing a proper acquisition function which can accurately model the behavior of the black-box function. Equipped with these probabilistic models, acquisition functions will be sequentially induced in order to leverage the uncertainty in the posterior for leading the exploration [34]. After observing the output value of each selected point in that iteration, the prior understanding of the black-box function and the acquisition function will be updated subsequently. Algorithm 1 illustrates the procedure of Bayesian optimisation.

But as we have mentioned before, many real-world optimisation problems are dealing with unknown constraints. In the next section we are investigating the role of unknown constraints in single-objective Bayesian optimisation.

Algorithm 1 Bayesian optimisation Algorithm

Require:
  1: **for** $n = 1, 2, ...,$ **do**
  2:      Optimise acquisition function α, $\mathbf{x}_{t+1} = \underset{\mathbf{x} \in \mathcal{X}}{\mathrm{argmax}} \; \alpha(\mathbf{x}; \mathcal{D}_n)$
  3:      Evaluate $\mathbf{x}_{t+1}$, obtain y_{t+1}
  4:      Augment data to the observation set, $\mathcal{D}_{n+1} = \{\mathcal{D}_n, (\mathbf{x}_{n+1}, y_{n+1})\}$
  5:      Update the prior model based on the new observed point.
  6: **end for**

2.1 Constrained single-objective Bayesian optimisation

Consider an SVM model; when tuning SVM hyperparameters, we may want to optimise performance on limited hardware (such as accessible memory). So, in addition to expensive evaluations of the objective function, we may face with similarly expensive evaluations of constraint functions. The only single-objective Bayesian optimisation with constraints method is proposed by [11]. In this paper, the authors have extended Bayesian optimization to incorporate inequality constraints. The aim of this method is:

$$\min_{c(\mathbf{x}) \leq \lambda} f(\mathbf{x}) \tag{2.1}$$

where $f(\mathbf{x})$ and $c(\mathbf{x})$ are both expensive and black-box functions. Authors defined a new *Constrained Improvement* in order to form the acquisition function. Constrained improvement is defined as:

$$\mathrm{I}_C(\hat{x}) = \Delta(\hat{x}) \; max\{0, f(x^+) - f(\hat{x})\} = \Delta(\hat{x})\mathrm{I}(\hat{x}) \tag{2.2}$$

In equation 3.4, $\Delta(\hat{x})$ is defined to be $\Delta(\hat{x}) \in \{0, 1\}$ which is a feasibility indicator function that returns 1 when $c(\hat{x}) \leq \lambda$ and 0 otherwise. Also x^+ denotes a feasible point with lowest function value observed in time τ. Due to dealing with black-box functions for both objectives and constraints, the authors use Bayesian formalism to model each with a Gaussian Process $\hat{c}(\mathbf{x}) \sim \mathcal{N}(\hat{\mu}_c(\mathbf{x}), \hat{\Sigma}_c(x))$ and $\hat{f}(\mathbf{x}) \sim \mathcal{N}(\hat{\mu}_f(\mathbf{x}), \hat{\Sigma}_f(x))$. Due to ehe marginal Gaussianity of $\hat{c}(\mathbf{x})$, the expected constrained improvement acquisition function is defined as:

$$\mathrm{EI}_C(\hat{x}) = \mathbb{E}[\mathrm{I}_C(\hat{x})|\hat{x}]$$
$$= \mathbb{E}[\Delta(\hat{x})\mathrm{I}(\hat{x})|\hat{x}]$$
$$= \mathbb{E}[\Delta(\hat{x})|\hat{x}]\mathbb{E}[\mathrm{I}(\hat{x})|\hat{x}] = \mathrm{PF}(\hat{x})\mathbb{E}[\mathrm{I}(\hat{x})$$

which the PF($\hat{\mathbf{x}}$) is defined as a simple univariate Gaussian cummulative distribution function PF($\hat{\mathbf{x}}$) $= Pr[c(\hat{\mathbf{x}}) \leq \lambda] = \int_{-\infty}^{\lambda} p(c(\hat{\mathbf{x}})|\hat{\mathbf{x}}, \tau_c) dc(\hat{\mathbf{x}})$. Thus the expected constrained improvement acquisition function $\text{EI}_C(\hat{\mathbf{x}})$ is the expected improvement of $\hat{x}$ over the best feasible point observed so far. It is also possible to extend the constraint function to the set of independant constraint functions $c(\mathbf{x}) = [c_1(\mathbf{x}), c_2(\mathbf{x}), ..., c_k(\mathbf{x})]$ [11]. In this case, PF($\hat{\mathbf{x}}$) is multipication of k constrains as $Pr[c(\mathbf{x})_i \leq \lambda_i]_{i=1...k}$. The authors evaluated their proposed model on one simulation function and two real-world problem of Locality Sensitive Hashing and SVM compression. The proposed method proved to be robust in the case of constraints existance.

2.2 Multi-objective optimisation and Bayesian optimisation

There are many studies on the topic of multi-objective optimisation. A quick summary of the most related studies have been illustrated in the Table 2.1.

Table 2.1: Summary of related studies in Multi-objective optimisation.

Study	Method	Constraints	Ranking
A Bayesian approach to constrained single- and multi-objective optimization [10]	**Bayesian optimization**	✓	-
Predictive Entropy Search for Multi-objective Bayesian Optimization with Constraints [14]	**Bayesian optimization**	✓	-
Predictive Entropy Search for Multi-objective Bayesian Optimization [17]	**Bayesian optimisation**	-	-
Pareto Frontier Learning with Expensive Correlated Objectives [33]	**Bayesian optimisation**	-	-
Pareto front modeling for sensitivity analysis in multi-objective bayesian optimization [5]	**Bayesian optimisation**	-	-

Active Learning of Pareto Fronts [6]	Active Learning for Regression Task; Gaussian Processes	✓	-
A Generative Kriging Surrogate Model for Constrained and Unconstrained Multi-objective Optimization [20]	Generative surrogate modeling	✓	-
Multi-objective Reinforcement Learning with Continuous Pareto Frontier Approximation [31]	gradient-based approach	✓	-
Active Learning for Multi-Objective Optimization [44]	Pareto Active Learning (PAL); Gaussian Processes	-	-
Faster Computation of Expected Hypervolume Improvement [19]	Expected Hypervolume Improvement	-	-
Multi-objective Bandits Optimizing the Generalized Gini Index [4]	Generalized Gini Index; gradient-based algorithm	-	-

The authors in [10] proposed a new Bayesian optimization approach to solve multi-objective optimization problems with non-linear constraints. The constraints are handled by extended domination rule. Also a new expected improvement formulation is proposed. In particular, the new formulation makes it possible to work around the problems in which no feasible solutions are available from the start. Sequential Monte Carlo sampling techniques are used in the process of computation and optimization of the new expected improvement criterion. The reason for using Sequential Monte Carlo sampling is due to no closed-form expression of expected improvement criterion. The contribution of this article is twofold. The first part of the contribution is about formulation of new sampling criterion that handles multiple objectives and non-linear constraints simultaneously. The second part of the contribution lies in the numerical methods employed to compute and optimize the sampling criterion [10]. There is another study about multi-objective optimisation with constraints based on Bayesian optimisation. The authors in [14] have described an information-based ap-

proach which can handle multiple objectives and several constraints. Motivated by the lack of Entropy-based methods, PESMOC is based on the Predictive Entropy Search for Multi-objective Bayesian Optimization [17]. At each iteration, PESMOC evaluates the objective functions and the constraints at an input location that is expected to reduce the entropy of the posterior distribution of the Pareto set the most. The proposed method is most useful useful in practical situations in which the objectives and the constraints are very expensive to evaluate [14]. There are studies about multi-objective optimisation. The study conducted in [17] described PESMO, a method for multi-objective Bayesian optimization. PESMO evaluates the objective functions at the input location that is most expected to reduce the entropy of posterior estimate of the Pareto set. The structure of acquisition function of PESMO can be understood as a sum of K individual acquisition functions, one per each of the K objectives. This triggers a decoupled evaluation scenario, in which the most promising objective is calculated by maximizing the individual acquisition functions. The other study in multi-objective Bayesian optimisation is proposed by [33]. The authors focus on modelling correlations amongst objectives in multi-objective Pareto optimization problems. To overcome the problem of intractable integrals in the proposed method, they have designed a novel approximation which leads to an analytic and differentiable approximation to the expected increase in Pareto hypervolume acquisition function. Another related study has been conducted in [5]. The authors showed that by computation of arbitrarily dense and continuous Pareto front, they can approximate the real Pareto front better in presence of measurement noise. These are useful tools to assist the user while making final decision among the Pareto points. An sudy based on Active Learning is proposed in [6]. Active Learning of Pareto Fronts framework adopts a different strategy. Pareto-optimal objective vectors are generated by combining the active learning paradigm with the solution of a scalarized optimization problem. The proposed model was iteratively refined until the information gain obtained by the new candidate training examples became negligible. A generative surrogate modeling procedure proposed in [20]. In this work, the authors have proposed a generative surrogate modeling procedure for multiobjective optimization. The main idea is *"finding a particular Pareto-optimal solution helps in modeling and finding another neighboring Pareto-optimal solution"*. The authors in [31] have proposed PMGA, a novel gradientbased approach to learn a continuous approximation of the Pareto frontier in MultiObjective Markov Decision Problems (MOMDPs). The idea is to define a parametric function that describes a manifold in the policyparameter space. The authors

have presented different alternatives, discussed about the advantages and disadvantages of the model and shown their properties through an empirical analysis. The presented model in [44], uses a Gaussian processes to predict the objective functions and to guide the sampling process in order to improve the prediction of the Pareto optimal set. it can be intuitively parameterized to achieve the desired level of accuracy at the lowest possible cost of evaluation. The authors presented an extensive theoretical analysis including bounds for the required number of evaluations to achieve the final accuracy. There is an study in which the authors use Gini index in the process of optimisation [4]. They introduced a new problem in the context of multiobjective multi-armed bandit (MOMAB). Contrary to most previously proposed approaches in MOMAB, They have tried to search for the Pareto front, instead we aim for a fair solution. To incorporate the fairness into the formulations, they have used the Generalized Gini Index (GGI), a well-known criterion developed in economics [4].

2.3 Multi-objective Bayesian optimisation with constraints

The goal of optimisation in such cases is to minimize the number of black-box function evaluations to find the global optimum of the function with respect to constraints. Bayesian optimisation with inequality constraints [11] and predictive entropy search for Bayesian optimization with unknown constraints [14] are two major studies investigating the role of inequality black-box expensive constraints in singleobjective Bayesian optimisation. There are also two recent studies on multi-objective Bayesian optimisation with constraints. For example, in [10] authors proposed a Bayesian multi-objective optimisation (BMOO) approach to solve the single-objective and multi-objective optimisation with non-linear constraints which is in the same context with the proposed problem in this paper. The method handles the constraints using an extended Pareto domination rule that takes both objectives and constraints into account. The authors evaluated their method on the benchmark test functions with respect to hypervolume improvement. They proposed approach is inspired from [29] which relies on highly complex data models. BMOO uses Sequential Monte-Carlo (SMC) in order to compute the integral over the expected improvement formulation. In [14], the authors proposed a method based on predictive entropy search. The authors generated 100 synthetic optimization problems obtained by sampling the objectives and the constraints from their respective GP prior and they did not use benchmark test functions to evaluate their method. This method is generally

categorized as informationbased methods while our proposed problem is based on hypervolume improvement approaches

2.4 Evolutionary algorithms and Multi-objective problems with constraints

Generally evolutionary algorithms and Surrogate-assisted evolutionary computation are not designed to work on limited budget of evaluation. But we are covering the overal review of the most related ones. Table 2.2 illustrates the related studies about such evolutionary methods.

Table 2.2: Summary of related studies in Evolutionary Multi-objective optimisation.

A surrogate-assisted evolution strategy for constrained multi-objective optimization [8]	Surrogate-assisted evolution	✓	-
A Simple and Fast Hypervolume Indicator-Based Multiobjective Evolutionary Algorithm [22]	Evolutionary Algorithm; Hypervolume	-	-
Multi-objective Evolutionary Algorithm for a Quick Computation of Pareto-Optimal Solutions [22]	Evolutionary Algorithm	✓	-
A fast and elitist multiobjective genetic algorithm: NSGA-II [9]	Evolutionary Algorithm	✓	-
SPEA2: Improving the strength Pareto evolutionary algorithm [43]	Evolutionary Algorithm	✓	-

A surrogate-assisted evolution strategy is proposed in [8]. They have developed a surrogate-assisted multi-objective evolution strategy (SMES) for computationally expensive constrained multiobjective optimization. The main limitation of the present work is that, as with most evolutionary algorithms, there is no theoretical guarantee that the proposed surrogate-assisted ES will converge to the Pareto front. Moreover, the proposed approximations could be inaccurate when the computational budget is severely limited due to the computational expense of the simulations. The authors in

[22] proposed a way for finding high quality of solutions in indicator-based multiobjective evolutionary algorithms (MOEAs). Hypervolume (HV) is a critical performance factor playing a role in solution selection. In this paper, a simple and fast hypervolume indicator-based MOEA (FV-MOEA) is proposed to quickly update exact HV contributions for different solutions. The core idea of FV-MOEA is that the HV contribution of a solution is only associated with parts of the solutions rather than the full solution set. The authors in [22] introduced a new method called ϵ-MOEA. They have proven that ϵ-MOEA has two main advantages: 1. It has helped in reducing the cardinality of Pareto-optimal region and and 2. It has also ensured that no two obtained solutions are within an ϵ_i from each other in the $i-th$ objective. Also NSGA-II [9] is one of the most famous evolutionary multi-objective algorithm. The authors have proposed a computationally fast and elitist MOEA based on a nondominated sorting approach. The authors believe the proposed method has less computational complexity of nondominated sorting, more elite, and lack of need for specifying the sharing parameter. The last study is about SPEA2 [43]. The authors name many advantages over the similar strategies:

- SPEA2 performs better than SPEA on all problem sets.

- PESA has fastest convergence power, probably due to its higher elitism intensity, but has difficulties on some problems regarding the boundry solutions because it does not always keep the boundary solutions.

- In higher dimensional objective spaces, SPEA2 seems to have advantages over NSGA-II.

2.5 Multi-objective problems with constraints and objective rankings

As we have already explained in section 1.1, there are many real-world problems, requiring handling the importance of the objectives. In other words, objective ranking. This informaion could either be used as priori or as an output of the algorithm. There are not any specific studies on this problem. However, there is an study [24] about the ranking of the solutions based on the objectives. The authors studied on an alternative dominance relation to Pareto-dominance. Based on each separate objective, ranking set of solutions will be calculated and an aggregation function is used to calculate a scalar fitness value for each solution. The authors called this relation

as *ranking-dominance* and it can be used to sort a set of solutions even for many objectives when Pareto-dominance relation is not able to distinguish solutions from one another.

Later in this proposal, we will explain our presented idea based on Bayesian optimisation for this particular problem and will present the corresponding results.

Chapter 3

Background

This chapter describes the background and related work on which this proposal builds. We will start from Gaussian process as a power tool for placing the prior belief and compute the uncertainty over the function space. Then we will the Bayesian optimisation and the acquisition functions. Finally we will explain about multi-objective Bayesian optimisation and Exptected Hypervolume Improvement.

3.1   Gaussian Processes

A Gaussian process is a collection of random variables $\{f(\mathbf{x}) : \mathbf{x} \in \mathcal{X}\}$ in which any finite collection of random variables has a multivariate Gaussian distribution [32]. GPs are determined by using a mean function $\mu(\mathbf{x}) : \mathcal{X} \to \mathbb{R}$ and a covariance function $k(\mathbf{x}) : \mathcal{X} \times \mathcal{X} \to \mathbb{R}^+$:

$$f(\mathbf{x}) \sim \mathcal{GP}(\mu(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')). \tag{3.1}$$

The particular choice of covariance function determines the properties of sample functions drawn from the GP prior. Popular kernels for covariance function are Squared Exponential, Matern, Periodic, and Linear [42]. Without loss of generality, the prior mean function generally assumed to be zero in Gaussian process. Let us assume that we already made t observations on the points $\mathbf{x}_{1..t} = \{\mathbf{x}_1, ..., \mathbf{x}_t\}$. The obtained evaluated results of these t observations are $f(\mathbf{x}_{1..t}) = \{f(\mathbf{x}_1), ..., f(\mathbf{x}_t)\}$. Based on the defined properties of Gaussian process, the function values $f(\mathbf{x}_{1..t})$ jointly follow a multivariate Gaussian distribution $f(\mathbf{x}_{1..t}) \sim \mathcal{N}(0, K)$ [25]. Where K is a positive definite kernel matrix:

$$K = \begin{bmatrix} k(\mathbf{x}_1, \mathbf{x}_1) & \dots & k(\mathbf{x}_1, \mathbf{x}_t) \\ \vdots & \ddots & \\ k(\mathbf{x}_t, \mathbf{x}_1) & \dots & k(\mathbf{x}_t, \mathbf{x}_t) \end{bmatrix}.$$

3.1.1 Computing the Posterior

The posterior can be similarly derived the way how the update equations for the Kalman filter was derived. First we need to find the joint distribution of $[f(\mathbf{x}_{t+1}), f(\mathbf{x}_1), f(\mathbf{x}_2), ..., f(\mathbf{x}_t)]$. Considering t observations of the objective function, for a new point $\mathbf{x}_{t+1}$, $P(f(\mathbf{x}_{t+1})|\mathbf{x}_{1..t}, f(\mathbf{x}_{1..t})) = \mathcal{N}(\mu_t(\mathbf{x}_{t+1}), \sigma_t^2(\mathbf{x}_{t+1}))$:

$$\begin{bmatrix} f(\mathbf{x}_{t+1}) \\ f(\mathbf{x}_1) \\ f(\mathbf{x}_2) \\ \vdots \\ f(\mathbf{x}_t) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} k(\mathbf{x}_{t+1}, \mathbf{x}_{t+1}) & k(\mathbf{x}_{t+1}, \mathbf{x})^T \\ k(\mathbf{x}_{t+1}, \mathbf{x}) & K_{\mathbf{xx}} \end{bmatrix} \right) \tag{3.2}$$

Where posterior mean and variance are calculated as:

$$\mu_t(\mathbf{x}_{t+1}) = k^T K^{-1} f(\mathbf{x}_{1..t}), \tag{3.3}$$

$$\sigma_t^2(\mathbf{x}_{t+1}) = k(\mathbf{x}_{t+1}, \mathbf{x}_{t+1}) - k^T K^{-1} k. \tag{3.4}$$

k in equation 3.3 and 3.4, is a kernel vector such that $k = [k(\mathbf{x}_{t+1}, \mathbf{x}_1), ..., k(\mathbf{x}_{t+1}, \mathbf{x}_t)]^T$. After modelling the observations with GP, it is required to find the best possible point for next iteration of optimisation ($\mathbf{x}_{t+1}$). In our proposed problem, the set of parameters that maximize the similarity of BurntOrangeexperimental results of BurntOrangethe Impact factors with the estimated function based on simulators.

Chapter 4

Sample Title

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Appendix A

Sample Title

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Appendix B

Sample Title

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Bibliography

[1] Majid Abdolshah, Alistair Shilton, Santu Rana, Sunil Gupta, and Svetha Venkatesh. Expected hypervolume improvement with constraints. In *Pattern Recognition (ICPR), 2018 25rd International Conference on*. IEEE, 2018.

[2] Eric Brochu, Tyson Brochu, and Nando de Freitas. A bayesian interactive optimization approach to procedural animation design. In *Proceedings of the 2010 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, pages 103–112. Eurographics Association, 2010.

[3] Eric Brochu, Vlad M Cora, and Nando De Freitas. A tutorial on bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. *arXiv preprint arXiv:1012.2599*, 2010.

[4] Robert Busa-Fekete, Balazs Szorenyi, Paul Weng, and Shie Mannor. Multiobjective bandits: Optimizing the generalized gini index. *arXiv preprint arXiv:1706.04933*, 2017.

[5] Roberto Calandra, Jan Peters, and MP Deisenrothy. Pareto front modeling for sensitivity analysis in multi-objective bayesian optimization. In *NIPS Workshop on Bayesian Optimization*, volume 5, 2014.

[6] Paolo Campigotto, Andrea Passerini, and Roberto Battiti. Active learning of pareto fronts. *IEEE transactions on neural networks and learning systems*, 25(3):506–519, 2014.

[7] Ivo Couckuyt, Dirk Deschrijver, and Tom Dhaene. Fast calculation of multiobjective probability of improvement and expected improvement criteria for pareto optimization. *Journal of Global Optimization*, 60(3):575–594, 2014.

[8] Rituparna Datta and Rommel G Regis. A surrogate-assisted evolution strategy for constrained multi-objective optimization. *Expert Systems with Applications*, 57:270–284, 2016.

[9] Kalyanmoy Deb, Amrit Pratap, Sameer Agarwal, and TAMT Meyarivan. A fast and elitist multiobjective genetic algorithm: Nsga-ii. *IEEE transactions on evolutionary computation*, 6(2):182–197, 2002.

[10] Paul Feliot, Julien Bect, and Emmanuel Vazquez. A bayesian approach to constrained single-and multi-objective optimization. *Journal of Global Optimization*, 67(1-2):97–133, 2017.

[11] Jacob R Gardner, Matt J Kusner, Zhixiang Eddie Xu, Kilian Q Weinberger, and John P Cunningham. Bayesian optimization with inequality constraints. In *ICML*, pages 937–945, 2014.

[12] Roman Garnett, Michael A Osborne, and Philipp Hennig. Active learning of linear embeddings for gaussian processes. *arXiv preprint arXiv:1310.6740*, 2013.

[13] Roman Garnett, Michael A Osborne, and Stephen J Roberts. Bayesian optimization for sensor set selection. In *Proceedings of the 9th ACM/IEEE International Conference on Information Processing in Sensor Networks*, pages 209–219. ACM, 2010.

[14] Eduardo C Garrido-Merchán and Daniel Hernández-Lobato. Predictive entropy search for multi-objective bayesian optimization with constraints. *arXiv preprint arXiv:1609.01051*, 2016.

[15] Michael A Gelbart, Jasper Snoek, and Ryan P Adams. Bayesian optimization with unknown constraints. *arXiv preprint arXiv:1403.5607*, 2014.

[16] Michael Adam Gelbart. *Constrained Bayesian Optimization and Applications*. PhD thesis, 2015.

[17] Daniel Hernández-Lobato, Jose Hernandez-Lobato, Amar Shah, and Ryan Adams. Predictive entropy search for multi-objective bayesian optimization. In *International Conference on Machine Learning*, pages 1492–1501, 2016.

[18] Matthew Hoffman, Bobak Shahriari, and Nando Freitas. On correlation and budget constraints in model-based bandit optimization with application to automatic machine learning. In *Artificial Intelligence and Statistics*, pages 365–374, 2014.

[19] Iris Hupkens, Michael Emmerich, and André Deutz. Faster computation of expected hypervolume improvement. *arXiv preprint arXiv:1408.7114*, 2014.

[20] Rayan Hussein and Kalyanmoy Deb. A generative kriging surrogate model for constrained and unconstrained multi-objective optimization. In *Proceedings of the Genetic and Evolutionary Computation Conference 2016*, pages 573–580. ACM, 2016.

[21] Frank Hutter, Holger H Hoos, and Kevin Leyton-Brown. Sequential model-based optimization for general algorithm configuration. In *International Conference on Learning and Intelligent Optimization*, pages 507–523. Springer, 2011.

[22] Siwei Jiang, Jie Zhang, Yew-Soon Ong, Allan N Zhang, and Puay Siew Tan. A simple and fast hypervolume indicator-based multiobjective evolutionary algorithm. *IEEE Transactions on Cybernetics*, 45(10):2202–2213, 2015.

[23] Tinu Theckel Joy, Santu Rana, Sunil Gupta, and Svetha Venkatesh. Hyperparameter tuning for big data using bayesian optimisation. In *Pattern Recognition (ICPR), 2016 23rd International Conference on*, pages 2574–2579. IEEE, 2016.

[24] Saku Kukkonen and Jouni Lampinen. Ranking-dominance and many-objective optimization. In *Evolutionary Computation, 2007. CEC 2007. IEEE Congress on*, pages 3983–3990. IEEE, 2007.

[25] Cheng Li, David Rubín de Celis Leal, Santu Rana, Sunil Gupta, Alessandra Sutti, Stewart Greenhill, Teo Slezak, Murray Height, and Svetha Venkatesh. Rapid bayesian optimisation for synthesis of short polymer fiber materials. *Scientific reports*, 7(1):5683, 2017.

[26] Daniel J Lizotte, Tao Wang, Michael H Bowling, and Dale Schuurmans. Automatic gait optimization with gaussian process regression. In *IJCAI*, volume 7, pages 944–949, 2007.

[27] Roman Marchant and Fabio Ramos. Bayesian optimisation for intelligent environmental monitoring. In *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, pages 2242–2249. IEEE, 2012.

[28] Ruben Martinez-Cantin, Nando de Freitas, Arnaud Doucet, and José A Castellanos. Active policy learning for robot planning and exploration under uncertainty. In *Robotics: Science and Systems*, volume 3, pages 334–341, 2007.

[29] Akira Oyama, Koji Shimoyama, and Kozo Fujii. New constraint-handling method for multi-objective and multi-constraint evolutionary optimization.

Transactions of the Japan Society for Aeronautical and Space Sciences, 50(167):56–62, 2007.

[30] Panos M Pardalos, Antanas Žilinskas, and Julius Žilinskas. *Non-convex multi-objective optimization*, volume 123. Springer, 2017.

[31] Matteo Pirotta, Simone Parisi, and Marcello Restelli. Multi-objective reinforcement learning with continuous pareto frontier approximation. In *29th AAAI Conference on Artificial Intelligence, AAAI 2015 and the 27th Innovative Applications of Artificial Intelligence Conference, IAAI 2015*, pages 2928–2934. AAAI Press, 2015.

[32] Carl Edward Rasmussen and Christopher KI Williams. *Gaussian processes for machine learning*, volume 1. MIT press Cambridge, 2006.

[33] Amar Shah and Zoubin Ghahramani. Pareto frontier learning with expensive correlated objectives. In *International Conference on Machine Learning*, pages 1919–1927, 2016.

[34] Bobak Shahriari, Kevin Swersky, Ziyu Wang, Ryan P Adams, and Nando De Freitas. Taking the human out of the loop: A review of bayesian optimization. *Proceedings of the IEEE*, 104(1):148–175, 2016.

[35] Jasper Snoek, Hugo Larochelle, and Ryan P Adams. Practical bayesian optimization of machine learning algorithms. In *Advances in neural information processing systems*, pages 2951–2959, 2012.

[36] Niranjan Srinivas, Andreas Krause, Sham M Kakade, and Matthias Seeger. Gaussian process optimization in the bandit setting: No regret and experimental design. *arXiv preprint arXiv:0912.3995*, 2009.

[37] Matthew Tesch, Jeff Schneider, and Howie Choset. Expensive multiobjective optimization for robotics. In *Robotics and Automation (ICRA), 2013 IEEE International Conference on*, pages 973–980. IEEE, 2013.

[38] Chris Thornton, Frank Hutter, Holger H Hoos, and Kevin Leyton-Brown. Autoweka: Combined selection and hyperparameter optimization of classification algorithms. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 847–855. ACM, 2013.

[39] Pratibha Vellanki, Santu Rana, Sunil Gupta, David Rubin, Alessandra Sutti, Thomas Dorin, Murray Height, Paul Sanders, and Svetha Venkatesh. Process-constrained batch bayesian optimisation. In *Advances in Neural Information Processing Systems*, pages 3417–3426, 2017.

[40] Ziyu Wang, Babak Shakibi, Lin Jin, and Nando de Freitas. Bayesian multi-scale optimistic optimization. In *Artificial Intelligence and Statistics (AISTATS)*, 2014.

[41] Ziyu Wang, Masrour Zoghi, Frank Hutter, David Matheson, Nando De Freitas, et al. Bayesian optimization in high dimensions via random embeddings. In *IJCAI*, pages 1778–1784, 2013.

[42] Andrew Wilson and Ryan Adams. Gaussian process kernels for pattern discovery and extrapolation. In *International Conference on Machine Learning*, pages 1067–1075, 2013.

[43] Eckart Zitzler, Marco Laumanns, and Lothar Thiele. Spea2: Improving the strength pareto evolutionary algorithm. *TIK-report*, 103, 2001.

[44] Marcela Zuluaga, Guillaume Sergent, Andreas Krause, and Markus Püschel. Active learning for multi-objective optimization. In *International Conference on Machine Learning*, pages 462–470, 2013.