

Constraint Handling in Efficient Global Optimization

Samineh Bagheri
Cologne University of Applied
Sciences (TH Köln)
Gummersbach, Germany
samineh.bagheri@th-koeln.de

Wolfgang Konen
Cologne University of Applied
Sciences (TH Köln)
Gummersbach, Germany
wolfgang.konen@th-koeln.de

Richard Allmendinger
University of Manchester
Manchester, UK
richard.allmendinger@manchester.
ac.uk

Jürgen Branke
University of Warwick
Coventry, UK
juergen.branke@wbs.ac.uk

Kalyanmoy Deb
Michigan State University
East Lansing, USA
kdeb@egr.msu.edu

Jonathan Fieldsend
University of Exeter
Exeter, UK
J.E.Fieldsend@exeter.ac.uk

Domenico Quagliarella
Italian Aerospace Research Centre
(CIRA)
Via Maiorise, snc
Capua, Italy 81043
d.quagliarella@cira.it

Karthik Sindhya
University of Jyväskylä
Jyväskylä, Finland
karthik.sindhya@jyu.fi

ABSTRACT

Real-world optimization problems are often subject to several constraints which are expensive to evaluate in terms of cost or time. Although a lot of effort is devoted to make use of surrogate models for expensive optimization tasks, not many strong surrogate-assisted algorithms can address the challenging constrained problems. Efficient Global Optimization (EGO) is a Kriging-based surrogate-assisted algorithm. It was originally proposed to address unconstrained problems and later was modified to solve constrained problems. However, these type of algorithms still suffer from several issues, mainly: (1) early stagnation, (2) problems with multiple active constraints and (3) frequent crashes. In this work, we introduce a new EGO-based algorithm which tries to overcome these common issues with Kriging optimization algorithms. We apply the proposed algorithm on problems with dimension $d \leq 4$ from the G-function suite [16] and on an airfoil shape example.

KEYWORDS

Constraint optimization, expensive optimization, surrogate models, Kriging, Gaussian processes, EGO

ACM Reference format:

Samineh Bagheri, Wolfgang Konen, Richard Allmendinger, Jürgen Branke, Kalyanmoy Deb, Jonathan Fieldsend, Domenico Quagliarella, and Karthik Sindhya. 2017. Constraint Handling in Efficient Global Optimization. In *Proceedings of GECCO '17, Berlin, Germany, July 15-19, 2017*, 8 pages. DOI: <http://dx.doi.org/10.1145/3071178.3071278>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](http://permissions.acm.org).

GECCO '17, Berlin, Germany

© 2017 ACM. 978-1-4503-4920-8/17/07...\$15.00

DOI: <http://dx.doi.org/10.1145/3071178.3071278>

1 INTRODUCTION

A constrained optimization problem (COP) can be defined as the minimization of an objective function (fitness function) f subject to inequality constraint function(s) g_1, \dots, g_m :

$$\begin{aligned} \text{Minimize} \quad & f(\vec{x}), \quad \vec{x} \in [\vec{a}, \vec{b}] \subset \mathbb{R}^d \\ \text{subject to} \quad & g_i(\vec{x}) \leq 0, \quad i = 1, 2, \dots, m \end{aligned} \quad (1)$$

where \vec{a} and \vec{b} define the lower and upper bounds of the search space (a hyperrectangle).

Real-world COPs are often very expensive to evaluate which means only a very limited number of function evaluations is allowed in practice. Giannakoglou [9] presents an overview on efficient optimization approaches and shows indicative examples from aerodynamics. Therefore, a proper optimizer for this sort of problems should be able to find optimal or near-optimal solutions with a low number of function evaluations. Many different surrogate-assisted techniques have been developed to tackle expensive COPs [3, 4, 20, 22, 28]. The main idea adopted by nearly all surrogate-assisted optimization algorithms is to use fast mathematical or statistical models for the optimization process and only evaluate a new solution on the expensive function when the model needs to be updated. Among the existing multitude of surrogate modeling approaches with various properties, Kriging [8, 15] appears to be one of the most attractive techniques. Kriging, also known as Gaussian Processes, offers a strong modeling tool which provides estimation of the prediction uncertainty in addition to the approximation of the function. Although in recent years many studies have been undertaken with different Kriging-based unconstrained and constrained optimizers, there are still limitations associated with the use of Kriging models. One of the issues that almost all Kriging-based strategies face is that frequent crashes occur during the optimization process. The crashes usually happen when a new solution is located very close to a former one. We find that adding a noise variance is an effective way to handle such stability issues.

Another point is that most of the existing Kriging-based constrained optimization strategies are evaluated only on simple 2-dimensional benchmark functions mostly with only one active constraint [7, 10, 25–27]. On the other hand, real-world COPs are often multi-constrained and are not limited to 2-dimensional problems. This motivates us to develop a new Kriging-based optimization algorithm which avoids crashes and is applicable on challenging COPs. We evaluate the proposed algorithm on all benchmarks with dimension $d \leq 4$ taken from the challenging G-function suite [16]. Throughout this work we try to answer the following research questions:

- (H1) Is it possible to modify existing Kriging-based optimization algorithms to handle challenging COPs with multiple *active* constraints?
- (H2) Is it possible to balance the exploration of feasible and infeasible infill points in a proper way?

1.1 Related Work

Most Kriging-based constrained optimization algorithms make use of Kriging's statistical property, expected improvement function [14, 17], for efficiently solving global COPs. We can categorize such algorithms into three main groups.

(a) The first group of algorithms transform a constrained problem into an unconstrained problem. Schonlau et al. [27] is an example of type (a) algorithm. This algorithm suggests to maximize the multiplication of the expected improvement of the objective function and the probability of feasibility, which are both statistical measures determined from Kriging models of objective and constraint functions. This algorithm fails if the number of active constraints are large or if the objective function is very steep or flat around the feasibility boundary. Algorithms which tend to maximize the penalized expected improvement [18, 25] also belong to the first group.

(b) Another approach to address COPs is to solve a constrained sub-problem. As an example, we can name a work from Sasena et al [26] in which the expected improvement is maximized subject to the approximation of the constraint functions. Audet et al. [2] maximize the expected improvement subject to the expected violation of each constraint.

(c) Methods in the third category transform the constrained problems to multi-objective unconstrained problems and then use multi-objective optimizers. These methods often consider the expected improvement of the fitness function as one objective and one or more statistical properties of the constraint functions as other objective(s) [7, 12, 19]. Although Durantin et al. [7] show that the type (c) algorithms perform better than the existing algorithms from type (a) and (b), we have to consider that solving a multi-objective problem is a complex task. An increase in problem dimension or number of active constraints makes such algorithms very time-consuming.

To the best of our knowledge, the current state-of-the-art for solving the G-problems is SACOBRA [3] (self adjusting COBRA). SACOBRA uses RBF surrogate models. SACOBRA is *not* a Kriging-based COP, it does not use the EI approach.

We propose in this paper a new type-(a) algorithm called SOCU (Surrogate-Assisted Optimization encompassing Constraints and Uncertainties).

2 METHODS

2.1 Kriging Surrogate Models

Kriging is a statistical modeling technique based on Gaussian processes. This algorithm approximates the function $f(x)$ with the surrogate model

$$Y = \mu + \epsilon(x), \quad (2)$$

where μ is the average of the stochastic process and the error term $\epsilon(x)$ is normally distributed with mean 0 and variance $\sigma^2(x)$. The estimation of μ and σ are the heart of Kriging modeling and described in more detail in the standard Kriging literature [14, 15].

2.2 Expected Improvement with Constraints

Efficient global optimization (EGO) is an algorithm developed by Jones et al. [14] for unconstrained optimization based on Kriging. The main idea of the EGO algorithm – originally introduced by Mokus et al. [17] – is to balance between exploration and exploitation by maximizing the expected improvement in Eq. (4) during a sequential optimization process:

$$\begin{aligned} EI(x) &= E[\max(f_{min} - Y, 0)] \\ &= (f_{min} - \mu(x))\Phi\left(\frac{f_{min} - \mu(x)}{\sigma(x)}\right) + \sigma(x)\varphi\left(\frac{f_{min} - \mu(x)}{\sigma(x)}\right), \end{aligned} \quad (3)$$

where the plugin f_{min} is the fitness value of the best found solution and Φ and φ are the cumulative and probability density function of the standard normal distribution. Schonlau et al. [27] extended the EGO algorithm to handle inequality constraints. Their algorithm maximizes the penalized expected improvement function EI_p shown in Eq. (4) which is the product of the expected improvement (now with plugin f_{min} being the best *feasible* fitness value) and the feasibility function $F(x)$

$$EI_p(x) = EI(x) \cdot F(x) = EI(x) \cdot \prod_{i=1}^m P(g_i(x) < 0), \quad (4)$$

where $P(g_i(x) < 0)$ is the probability of $g_i(x)$ to be feasible, measured with the help of the Kriging model for the i th constraint:

$$P(g_i(x) < 0) = \Phi\left(\frac{-\mu_i(x)}{\sigma_i(x)}\right), \quad (5)$$

This algorithm often faces difficulties in solving COPs with 2 or more active constraints, because the product of feasibility probabilities approaches zero near the feasibility border where the optimum is located. Therefore, the penalized expected improvement may have very small values in the interesting region. Furthermore, this algorithm is unlikely to sample infeasible solutions. Others have shown that existence of the infeasible solutions in the population can often be helpful [23, 24]. In order to give solutions around the boundary a higher chance to be selected, we modify the feasibility function introduced by Schonlau [27] and formulate a modified expected improvement function EI_{mod} as follows:

$$EI_{mod}(x) = EI(x) \cdot F(x) = EI(x) \cdot \prod_{i=1}^m \min\left(2\Phi\left(\frac{-\mu_i(x)}{\sigma_i(x)}\right), 1\right) \quad (6)$$

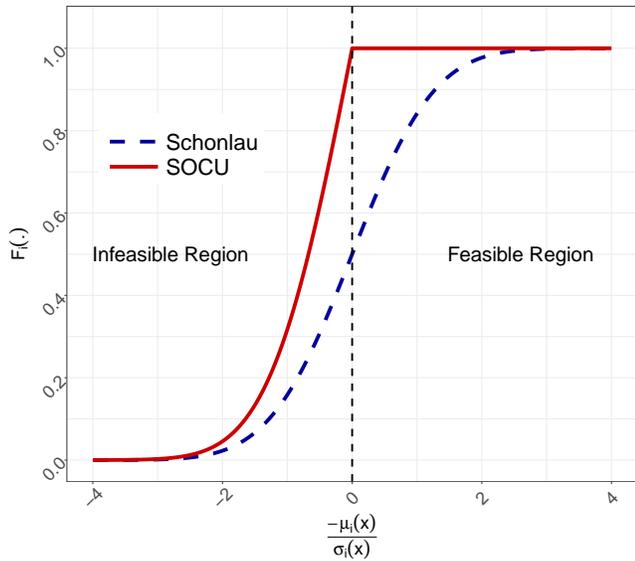


Figure 1: Feasibility function for the i -th constraint $F_i(x)$. The total feasibility function is $F(x) = \prod F_i(x)$

Fig. 1 shows the different feasibility functions used in Schonlau algorithm (blue dashed line) and our proposed algorithm (red line) for one constraint.

The proposed algorithm SOCU, shown in Alg. 1, initially maximizes the feasibility function in order to find at least one feasible solution. As soon as one feasible solution is found, the algorithm proceeds by maximizing the modified expected improvement (Eq. 6) in each iteration. Since the EI_{mod} function is highly multimodal, we decided to use a simulated annealing method as the internal optimizer.

2.3 Plugin Control (PC): Preserving Feasibility

In our first experiments with EI_{mod} we observed a strange behavior depicted in Fig. 2: Initially, the surface plot of EI_{mod} looks as expected (upper plot), but the best feasible solution is still far away from the true optimum. When finding better solutions near the true optimum, the EI_{mod} surface would suddenly change (lower plot) and the maximum of EI_{mod} shifts far away into infeasible area. What is the reason? A closer analysis revealed that the plugin f_{min} in Eq. (4) is responsible for this. Usually, the plugin f_{min} is taken as the best feasible objective found so far. A new value for f_{min} does not change the maxima locations of EI , but it changes the intercept. This has a large effect on the maxima locations of $EI_{mod}(x) = EI(x) \cdot F(x)$. We explain this with a 1D example.

2.3.1 **The 1D-Case.** Consider the following simple 1D-model:

$$EI(x) = \max(ax + b, 0) \tag{7}$$

$$F(x) = \min(2\Phi(kx), 1) \tag{8}$$

We assume $k > 0$, so that $x < 0$ is the infeasible area, and $a < 0$, i. e. the (unconstrained) $EI(x)$ has better values towards $x < 0$. This makes the constraint active, meaning that the constrained optimum is on the border $x = 0$.

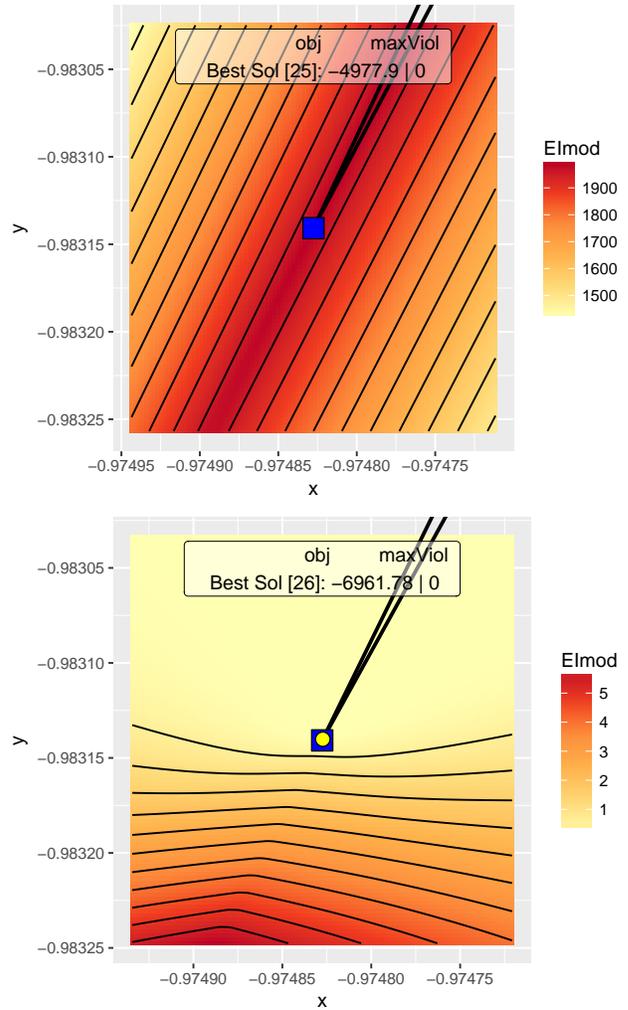


Figure 2: EI_{mod} shift towards the infeasible area for a 2-dimensional test problem (G06). The thick lines in form of a pointed triangle show the feasible area, the blue square is the true solution. The yellow circle is the best feasible solutions (being outside the plot area in the upper plot).

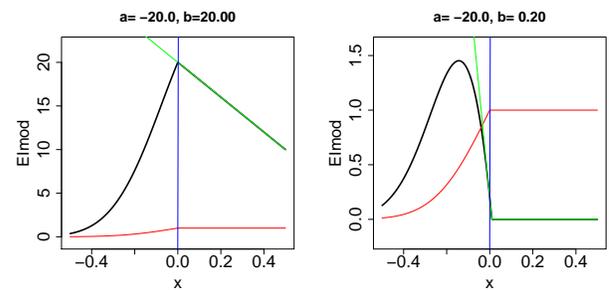


Figure 3: EI (green), F (red), and EI_{mod} (black) in the 1D-case. Right: For small intercepts b the optimum of EI_{mod} is shifted towards the infeasible area $x < 0$.

Algorithm 1 SOCU algorithm

```

1:  $m$ : number of constraints
2:  $n$ : number of evaluated points
3:  $d$ : dimension of the problem
4:  $pop^{(n)}$ : population of  $n = 5d$  initial points generated by LHS
5: while  $n \leq Budget$  do
6:   Build from  $pop^{(n)}$  the Kriging models for objective function
    $f: (\mu_0, \sigma_0)$  and the  $m$  constraints  $g_i: (\mu_1, \sigma_1), \dots, (\mu_m, \sigma_m)$ 
7:   Obtain  $EI(x)$  from Eq. (4) with plugin corrector Eq. (10)
8:    $F(x) = \prod_{i=1}^m \min\left(2\Phi\left(-\frac{\mu_i(x)}{\sigma_i(x)}\right), 1\right)$ 
9:    $EI_{mod}(x) = EI(x) \cdot F(x)$ 
10:  if a feasible solution has been found then
11:     $x_{new} = \arg \max(EI_{mod}(x))$   $\triangleright$  Use simulated annealing
12:  else
13:     $x_{new} = \arg \max(F(x))$ 
14:  end if
15:  Add  $x_{new}$  to  $pop^{(n)}$  and evaluate it on true  $f$  and  $g_1, \dots, g_m$ 
16:   $n \leftarrow n + 1$ 
17: end while
    
```

What happens now for EI_{mod} as a function of the intercept b ? As Fig. 3 depicts, large b have the optimum for EI_{mod} correctly at $x = 0$, but too small values for b lead to a false shift of EI_{mod} 's maximum towards the infeasible area. A short Taylor expansion shows that the critical intercept is

$$b_{crit} = \frac{a\sqrt{2\pi}}{2k}. \quad (9)$$

Smaller intercepts have the maximum of EI_{mod} shifted to the infeasible area.

To correct this, we simply have to change the plugin in Eq. (4):

$$EI(x) = E[\max(b_{crit} + f_{min} - Y, 0)] \quad (10)$$

This ensures that at the active border, where Y is not larger than f_{min} , the value of $EI(x)$ is at least b_{crit} . – If we have multiple constraints, we calculate b_{crit} for each of them

2.3.2 The 2D- and nD-Case. In the higher-dimensional case ($d > 1$) we have to find the direction \vec{g} of slowest descent of EI_{mod} at the current best feasible point. This is for example in the case of G06 the bisecting line of the two constraints. The slopes a and k for EI and the constraint(s) in Eq. (9) have to be replaced by the respective slopes along direction \vec{g} .

3 EXPERIMENTAL SETUP

3.1 General Setup

Initially we test the proposed algorithm on a toy problem *Sphere4* of steerable difficulty: This problem has a sphere as objective function and 4 linear constraints, 2 of them being active, 2 inactive. The constraints enclose a feasible region with a triangular tip of angle ϕ (see Fig. 4).

Next, we apply SOCU to all G-problems from the G-problem suite having 4 or less dimensions (see Table 1).¹ This is because it is well known that Kriging algorithms are viable only for not too

large dimensions. Equality constraints are translated to inequality constraints.²

For each algorithm we run 30 independent trials with different $n = 5d$ initial points. Constraint violations smaller than 10^{-5} are tolerated. We consider a fixed budget of 100 function evaluations for all problems. The Kriging models are built by R packages DICEKRIGING and DICEOPTIM. In order to optimize EI_{mod} we use Generalized Simulated Annealing (R package GENSA). The time limit for this internal optimizer is set to 10 seconds in each iteration, allowing for several thousand Kriging model evaluations (depending on problem size).

Additionally, we run SOCU on an application example from aerodynamics described in the next section.

3.2 Application example to an aerodynamic shape design problem

The benchmarks related to aerodynamic shape design problems represent an ideal platform to test optimization systems and algorithms that make use of surrogate methods for the evaluation of the objective functions and the constraints [13]. This is for two main reasons, namely that the evaluation of objectives and constraints very often requires a significant computational effort, and that both objectives and constraints may have a degree of non-linearity which is a function of the computational model used, the operating conditions considered and the shape parameterization chosen. The problem presented here is a simple subsonic airfoil section design exercise derived from [21], but with some special features in terms of ease of implementation, flexibility, reproducibility and availability of the analysis codes that allow its use even in contexts not specialized to aerodynamic design.

3.2.1 Problem setup. The goal of this optimization problem is the reduction of the aerodynamic drag of a given airfoil changing its shape. A generic airfoil shape is parameterized as linear combination of an initial geometry, defined parametrically by $(x_0(s), y_0(s))$, and a number of functions $y_i(s)$ that may be defined analytically or by point distributions [11]:

$$y(s) = k \left(y_0(s) + \sum_{i=1}^n w_i y_i(s) \right), \quad x(s) = x_0(s) \quad (11)$$

where the airfoil shape is controlled by the design parameters w_i and by the scale factor k . The operating conditions are Mach number equal to 0.0 (incompressible flow) and Reynolds number equal to 3000000. The starting airfoil is the NACA 2412 [1]. The Mach and Reynolds numbers that characterize the specified regime of operation are sufficiently low to allow an extended laminar bucket that can have a beneficial effect on a large part of the flight envelope. The design goals are translated into the following optimization problem:

$$\left\{ \begin{array}{l} \min C_D \\ \text{subject to: } C_L = 0.5 \\ C_M \geq -0.07 \\ C_{Dp} \geq 0 \\ t/c = 0.12 \\ \ell_R \geq 0.006 \end{array} \right. \quad (12)$$

¹G02 and G03 are problems scalable in their dimension d . We use here $d = 2$.

²The inequality sign is chosen in such a way that the constrained optimum stays at the same location.

Table 1: Characteristics of the G-functions: *d*: dimension, ρ^* : feasibility rate (%), *FR*: range of the fitness values, *GR*: ratio of largest to smallest constraint range, *LI/NI*: number of linear/nonlinear inequalities, *LE/NE*: number of linear/nonlinear equalities, *a*: number of constraints active at the optimum.

Fct.	<i>d</i>	ρ^*	<i>FR</i>	<i>GR</i>	<i>LI/NI</i>	<i>LE/NE</i>	<i>a</i>
G02mod	2	99.997%	0.57	2.632	1 / 1	0 / 0	1
G03mod	2	78.388%	1.99	1.000	0 / 0	0 / 1	1
G05	4	0.0000%	8863.69	1788.74	2 / 0	0 / 3	3
G06	2	0.0072%	1246828.23	1.010	0 / 2	0 / 0	2
G08	2	0.8751%	1821.61	2.393	0 / 2	0 / 0	0
G11	2	0.0000%	4.99	1.000	0 / 0	0 / 1	1
G15	3	0.0000%	586.0	1.034	0 / 0	1 / 1	2
G24	2	0.44250%	6.97	1.82	0 / 2	0 / 0	2
XFOIL	4	0.1349 %	0.99	1.34	0 / 3	0 / 0	1

Where C_D , C_{Dp} , C_L and C_M are the drag, pressure drag, lift and pitching moment coefficient of the airfoil; t/c denotes the thickness to chord ratio. The two equality constraints defined in (12) are here satisfied by explicitly changing two free problem parameters and therefore they are not considered by the optimization algorithm. In particular, the constraint on t/c is satisfied by changing properly the free parameter k , while the constraint on C_L is satisfied by changing the second free parameter, namely the airfoil angle of attack α .

The aerodynamic analysis code here selected to evaluate the airfoil performance is Drela’s XFOIL code [5]. This code is based on a second order panel method interactively coupled to a boundary layer integral module. Laminar to turbulent flow transition is predicted using the method described in [6].

4 RESULTS

Demonstration on *Sphere4*

The *Sphere4* problem is a constrained optimization problem of steerable difficulty. It is used here to demonstrate the difference between the Kriging algorithms. Fig. 5 shows the results on *Sphere4*. As expected, the problem gets harder for all algorithms as the feasibility angle ϕ decreases. This is because the feasible region becomes smaller and the optimum is surrounded by a neighborhood containing more and more infeasible points. – Interestingly, also the gap between Schonlau and SOCU gets larger as ϕ decreases. This is understandable as well: For SOCU the probability of feasibility is 1 along both active constraint lines for all ϕ . For Schonlau the probability of feasibility is 0.25 at the optimum (2 active constraints, both 0.5). If ϕ is large, the probability quickly rises to 0.5 as we move along one of the active constraint lines, because the distance to the other line increases. But if ϕ is small and we move along one constraint line, the probability according to Schonlau stays longer near 0.25, because the second constraint line is not far away. Thus the solution found by Schonlau will be farther from the optimum since the maximum of the product $EI \cdot F$ moves farther into the feasible region. This is exactly what we see in Fig. 5.

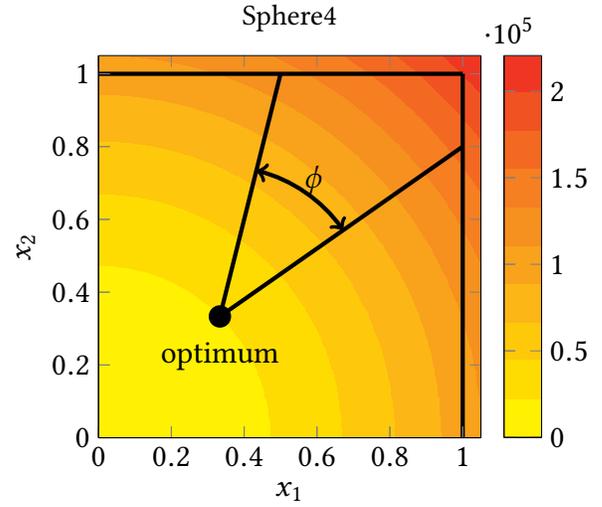


Figure 4: *Sphere4* problem. The colored contour levels show the objective function $10^5 \cdot (x_1^2 + x_2^2)$ (2D sphere function). The thick black lines depict 4 linear constraints enclosing the feasible region. The difficulty of this problem is scalable by changing the feasibility angle ϕ .

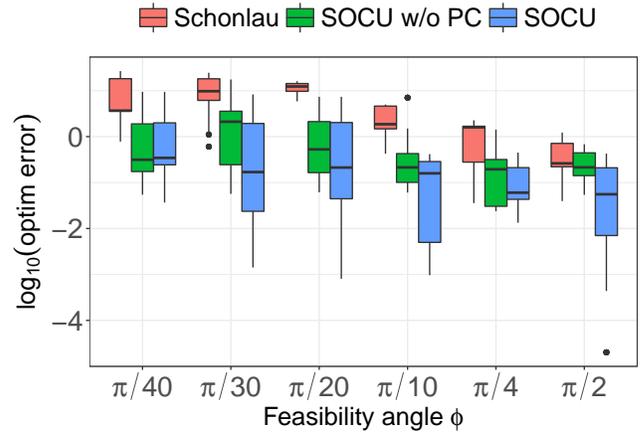


Figure 5: Comparing the final optimization error determined by different algorithms for optimizing *Sphere4* problems with different feasibility angles. The results are taken from 30 independent runs and 100 function evaluations.

Noise Variance

In our first experiments we experienced frequent crashes of the Kriging modeling software due to numeric instabilities. This is a well-known but cumbersome observation shared by many researchers in Kriging especially if the modeling points are unevenly spaced, as it is inevitably the case in optimization tasks. To avoid too strong oscillations of the Kriging model due to nearby points, it is a common cure to switch from interpolating to approximating Kriging models, either with the so-called nugget-effect or with a

Table 2: Effect of noise variance.

SOCU w/o noise variance			SOCU	
problem	crashed (%)	iteration	crashed (%)	iteration
G06	100	19	0	-
G02mod	96	40	0	-

noise variance parameter, which assumes a certain noise or uncertainty related with every modeling point. Since the nugget effect leads to a complicated structure for the variance $\sigma^2(x)$, it turned out to be not well-suited for our case.

The noise variance, on the other hand, turned out to be very effective: As Table 2 shows, the interpolating Kriging models had frequent crashes. By adding the noise variance $v = 0.01$ to the model, we could completely avoid any crashes in our experiments.

Performance on G-Problems

In Fig. 6 we compare the two different variants of Kriging-based, EGO, namely the original version of Schonlau et al. [27] and our SOCU algorithm, which we applied to all G-problems in Tab. 1. SOCU reaches lower optimization errors in most cases. In the case of G08, both algorithms have the same median curve. This is perfectly understandable, since G08 is the only problem without any active constraints. Absence of an active constraint is a convincing reason for similar performance of both algorithms, since the different feasibility functions (Fig. 1) have no effect. For problems with active constraints, the high value of SOCU's $F_i(x)$ at the border of the feasible region helps to find better solutions.

In Fig. 7 we show additionally the effect of switching off the plugin control (Sec. 2.3) in SOCU (green boxplot 'SOCU w/o PC'). It can be seen that the plugin control is beneficial for G03, G06 and G11 (blue boxplot 'SOCU').

We compare in Fig. 8 the results of SOCU with SACOBRA [3], to the best of our knowledge the current state-of-the-art for solving the G-problems. It is evident that SACOBRA is slightly better in most cases, except for the case of G24. Additionally, three problems (G05, G15, G24) show a better performance of SOCU in the early stages (between iteration 25 and 75), although SACOBRA catches up at iteration 100.

Finally Fig. 9 shows the overall performance comparison for all algorithms on all test problems in form of a data profile. The larger the data profile (ratio of solved problems), the better the relevant algorithm. The performance factor α on the x-axis is the number of iterations divided by $d + 1$.

Performance on XFOIL

Fig. 10 shows our results for the XFOIL case. SOCU is slightly better than Schonlau in the median, but the difference is not statistically significant. The similarity is understandable, since problem XFOIL has only one active constraint. In such a case we do expect the differences between Schonlau and SOCU to be not very large. SOCU and SACOBRA produce nearly identical results.

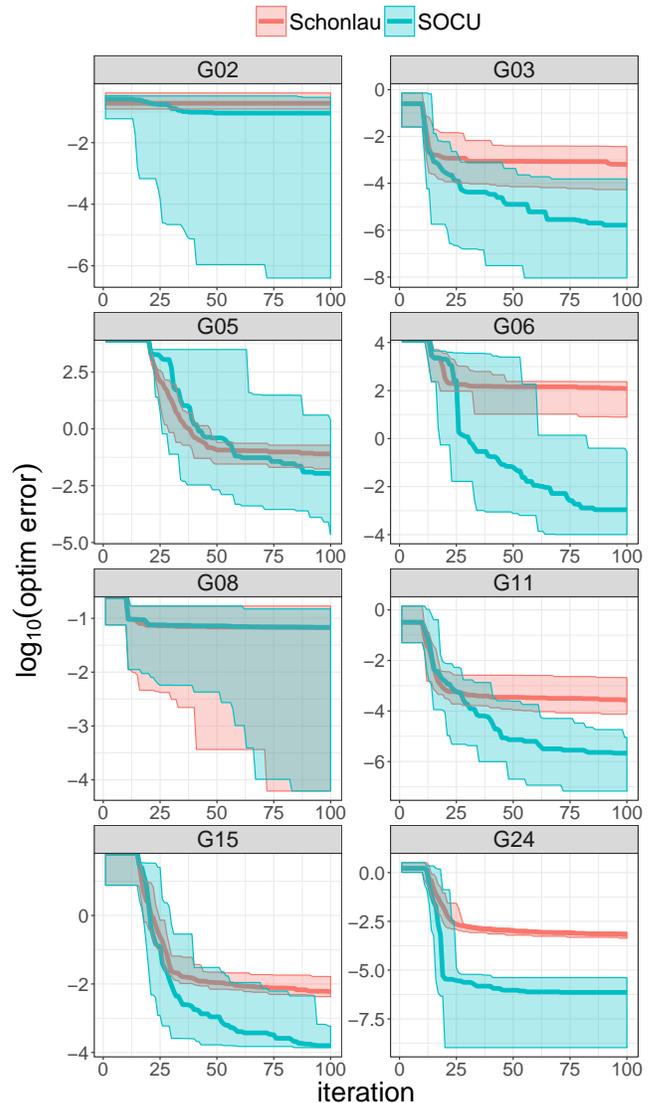


Figure 6: Comparing the performance of SOCU and Schonlau [27]. The thick lines show the median from 30 independent runs while the colored bands mark the worst and the best run.

5 DISCUSSION AND CONCLUSION

We applied Kriging surrogate models to constrained optimization. We could show that a small number of evaluations is sufficient to obtain good optimization results. We were not the first to do so, but three important conclusions for Kriging-based optimization could be drawn in this paper:

- (1) Interpolating Kriging models often suffer from numerical instabilities and subsequent crashes, especially when the population points are unevenly distributed in the search space. This will be nearly always the case when applying Kriging for optimization. We have shown that these

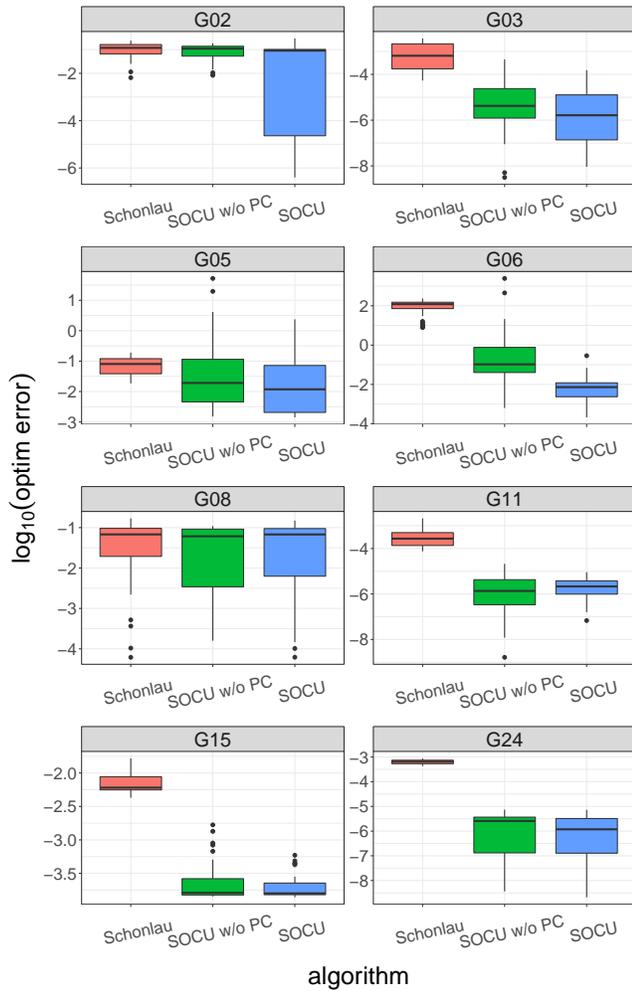


Figure 7: Comparing the final optimization error from 30 independent runs determined with different algorithms

crashes can be completely avoided (at least in our test cases) when we add a small noise variance to the Kriging models. This leads to approximating Kriging models and to a variance always larger than zero. Both effects are beneficial for numeric stability of the Kriging models.

- (2) Many Kriging-based COP-solvers use in one form or the other a product of expected improvement EI and probability of feasibility F . We could show that if the additive plugin in EI gets very small (which is no problem for the optima of EI themselves), this can have adversarial effects for the optima of $EI \cdot F$. We proposed a method called *plugin control* which successfully counteracts such adversarial effects.
- (3) Having this plugin control in effect, we could show that a probability curve of SOCU being 1 at the border of feasibility is clearly superior to an approach where the probability is 0.5 at the border [27]. The benefits are - as expected -

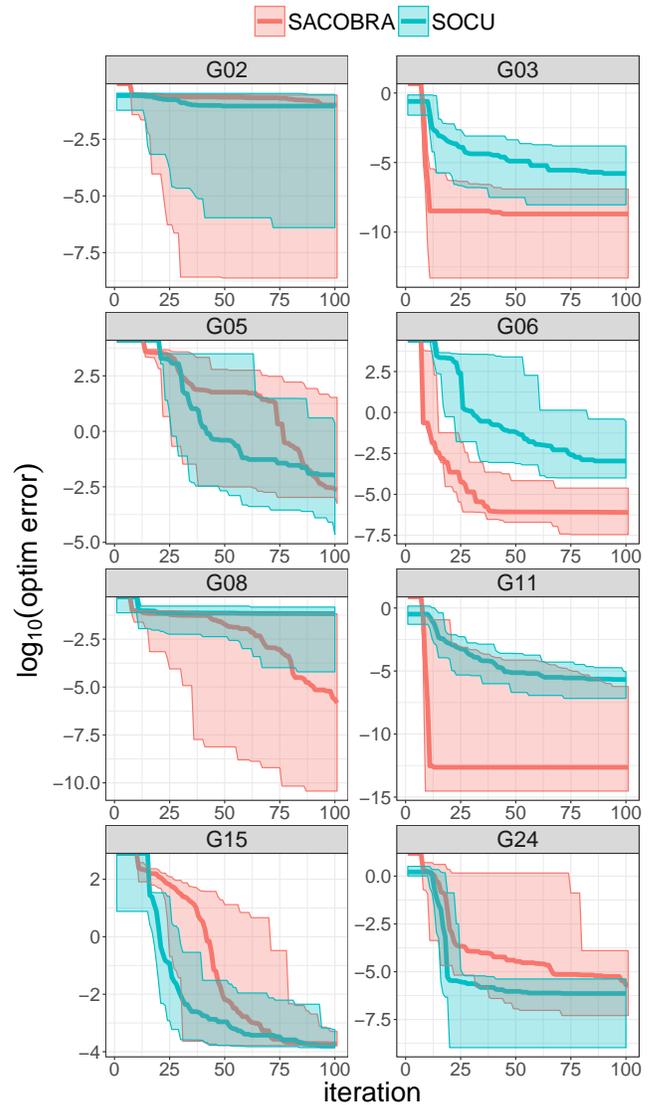


Figure 8: Comparing optimization performance of SACOBRA and SOCU.

more clearly seen for problems with two or more active constraints (G05, G06, G15, G24).

We tested our new algorithm SOCU on a variety of benchmark problems with dimension of $d = 4$ and below. We could perform better than the Kriging-based algorithm of Schonlau et al [27], but were in most cases worse than the non-Kriging-based SACOBRA algorithm [3].

In some cases (G05, G15, G24) SOCU showed a better performance (median, best and worst case) in early iterations (< 75) than SACOBRA. We suppose that this is due to the better exploration of the EGO approach. A drawback for Kriging-based models is however that they become very slow for larger dimensions.

In the future, a closer investigation into the causes of the observed performance differences would be desirable. Furthermore,

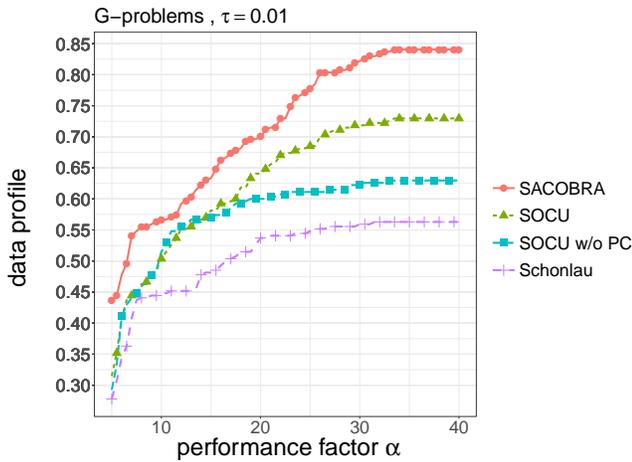


Figure 9: Data profile.

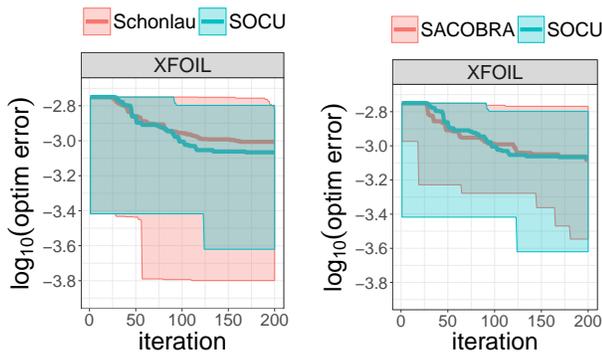


Figure 10: Comparing the performance of SOCU, Schonlau [27], and SACOBRA [3] on the XFOIL real world optimization problem (30 independent runs).

as the name SOCU suggests, we plan to test our algorithm also on problems involving uncertainty.

The source code of the proposed algorithm "SOCU" is available online³.

ACKNOWLEDGMENTS

This work was initiated during the Surrogate-Assisted Multi-Criteria Optimization Workshop at the Lorentz Center (Leiden, The Netherlands), 2016. We are grateful to the other participants of the workshop and the Lorentz Center for their support. This research was partly funded by Tekes, the Finnish Funding Agency for Innovation (the DeCoMo project), and by the Engineering and Physical Sciences Research Council [grant numbers EP/N017195/1, EP/N017846/1].

REFERENCES

[1] I. H. A. Abbot and A. E. von Doenhoff. 1959. *Theory of wing sections, including a summary of airfoil data*. Dover Publications, New York.

³<https://github.com/saminehbagheri/SOCU>

[2] C. Audet, A. J. Booker, Dennis, Jr, P. D. Frank, and D. W. Moore. 2000. A Surrogate-Model-Based Method For Constrained Optimization. In *AIAA/ISSMO*. 2000-4891.

[3] S. Bagheri, W. Konen, and T. Bäck. 2016. Online selection of surrogate models for constrained black-box optimization. In *IEEE SSCT2016*. 1-8.

[4] A. J. Booker, J. E. Dennis, P. D. Frank, D. B. Serafini, V. Torczon, and M. W. Trosset. 1999. A rigorous framework for optimization of expensive functions by surrogates. *Structural Optimization* 17, 1 (1999), 1-13.

[5] M. Drela. 1989. XFOIL: An Analysis and Design System for Low Reynolds Number Airfoils. In *Conference on Low Reynolds Number Airfoil Aerodynamics*. University of Notre Dame. http://web.mit.edu/drela/Public/papers/xfoil_sv.pdf

[6] M. Drela and M. B. Giles. 1987. Viscous-Inviscid Analysis of Transonic and Low Reynolds Number Airfoils. *AIAA Journal* 25, 10 (Oct. 1987), 1347-1355.

[7] C. Durantin, J. Marzat, and M. Balesdent. 2016. Analysis of multi-objective Kriging-based methods for constrained global optimization. *Computational Optimization and Applications* 63, 3 (2016), 903-926.

[8] A. I. Forrester and A. J. Keane. 2009. Recent advances in surrogate-based optimization. *Progress in Aerospace Sciences* 45, 1 (2009), 50-79.

[9] K.C. Giannakoglou. 2002. Design of optimal aerodynamic shapes using stochastic optimization methods and computational intelligence. *Progress in Aerospace Sciences* 38, 1 (2002), 43 - 76.

[10] R. B. Gramacy and H. K. H. Lee. 2011. Optimization under unknown constraints. In *Bayesian Statistics*. Vol. 9. 229-247.

[11] R. Hicks and P. A. Henne. 1978. Wing Design by Numerical Optimization. *Journal of Aircraft* 15, 7 (1978), 407-412.

[12] R. Hussein and K. Deb. 2016. A Generative Kriging Surrogate Model for Constrained and Unconstrained Multi-objective Optimization. In *GECCO '16*. ACM, New York, NY, USA, 573-580.

[13] E. Iuliano and D. Quagliarella. 2015. Evolutionary Optimization of Benchmark Aerodynamic Cases using Physics-based Surrogate Models. In *AIAA SciTech*. American Institute of Aeronautics and Astronautics, 1721-1736.

[14] D. R. Jones, M. Schonlau, and W. J. Welch. 1998. Efficient Global Optimization of Expensive Black-Box Functions. *J. of Global Optimization* 13, 4 (Dec. 1998), 455-492.

[15] D. G. Krige. 1951. A statistical approach to some basic mine valuation problems on the Witwatersrand. *Journal of the Southern African Institute of Mining and Metallurgy* 52, 6 (1951), 119-139.

[16] J. Liang, T. P. Runarsson, E. Mezura-Montes, M. Clerc, P. Suganthan, C. Coello, and K. Deb. 2006. Problem definitions and evaluation criteria for the CEC 2006 special session on constrained real-parameter optimization. *Journal of Applied Mechanics* 41, 8 (2006).

[17] J. Mockus. 1977. On Bayesian Methods for Seeking the Extremum and their Application. In *IFIP Congress*. 195-200.

[18] J. Parr, C. M. Holden, A. I. Forrester, and A. J. Keane. 2010. Review of efficient surrogate infill sampling criteria with constraint handling. In *2nd International Conference on Engineering Optimization*. 1-10.

[19] J. M. Parr, A. J. Keane, A. I. Forrester, and C. M. Holden. 2012. Infill sampling criteria for surrogate-based optimization with constraint handling. *Engineering Optimization* 44, 10 (2012), 1147-1166.

[20] Victor Picheny. 2014. A stepwise uncertainty reduction approach to constrained global optimization.. In *International Conference on Artificial Intelligence and Statistics*. Reykjavik, Iceland, 787-795.

[21] D. Quagliarella, G. Petrone, and G. Iaccarino. 2014. Optimization Under Uncertainty Using the Generalized Inverse Distribution Function. In *AISTATS*, W. Fitzgibbon (Ed.). Computational Methods in Applied Sciences, Vol. 34. Springer, NL, 171-190.

[22] R. G. Regis. 2014. Constrained optimization by radial basis function interpolation for high-dimensional expensive black-box problems with infeasible initial points. *Engineering Optimization* 46, 2 (2014), 218-243.

[23] T. P. Runarsson and X. Yao. 2000. Stochastic ranking for constrained evolutionary optimization. *IEEE Transactions on Evolutionary Computation* 4, 3 (2000), 284-294.

[24] T. P. Runarsson and X. Yao. 2005. Search biases in constrained evolutionary optimization. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews* 35, 2 (2005), 233-243.

[25] M. J. Sasena, P. Papalambros, and P. Goovaerts. 2002. Exploration of meta-modeling sampling criteria for constrained global optimization. *Engineering optimization* 34, 3 (2002), 263-278.

[26] M. J. Sasena, P. Y. Papalambros, and P. Goovaerts. 2001. The Use of Surrogate Modeling Algorithms to Exploit Disparities in Function Computation Time within Simulation-Based Optimization. In *4th World Congress of Structural and Multidisciplinary Optimization*. 5-11.

[27] M. Schonlau, W. J. Welch, and D. R. Jones. 1998. *Global versus local search in constrained optimization of computer models*. Lecture Notes-Monograph Series, Vol. 34. Institute of Mathematical Statistics, Hayward, CA, 11-25.

[28] D. Villanueva, R. Le Riche, G. Picard, and R. Haftka. 2012. Surrogate-based agents for constrained optimization. In *AIAA Non-Deterministic Approaches Conference*. 1935-1951.