A Scalarized Augmented Lagrangian Algorithm (SCAL) for Multi-objective Optimization Constrained Problems

Lino Costa[1,2], Isabel Espírito Santo[1,2] and Pedro Oliveira[3,4]
[1]Algoritmi R&D Center, University of Minho, Portugal
[2]Department of Production and Systems Engineering, University of Minho, Portugal
[3]Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto, Portugal
[4]ISPUP-EPIUnit, Universidade do Porto, Rua das Taipas, n° 135, 4050-600 Porto, Portugal

Keywords: Multi-objective Constrained Optimization, Augmented Weighted Tchebycheff, Pattern Search, Augmented Lagrangian.

Abstract: In this paper, a methodology to solve constrained multi-objective problems is presented, using an Augmented Lagrangian technique to deal with the constraints and the Augmented Weighted Tchebycheff method to tackle the multi-objective problem and find the Pareto Frontier. We present the algorithm, as well as some preliminary results that seem very promising when compared to previous state-of-the- art work. As far as we know, the idea of incorporating an Augmented Lagrangian in multi-objective optimization is rarely used so, the obtained results are very encouraging to pursuit further in this line of investigation, namely with the tuning of the Augmented Lagrangian parameters as well as testing other algorithms to solve the subproblems or to handle the multi-objective problems. It is also our intention to investigate the resolution of problems with three or more objectives.

1 INTRODUCTION

A multi-objective optimization (MO) problem with m objectives and n decision variables, without loss of generality, can be mathematically formulated as follows:

$$\begin{array}{ll} \text{minimize:} & f(x) = (f_1(x), f_2(x), \ldots, f_m(x))^{\mathrm{T}} \\ \text{subject to:} & \\ & c_i(x) = 0, \quad i = 1, \ldots, q \\ & g_j(x) \geq 0, \quad j = 1, \ldots, p \\ & x \in \Omega \end{array} \tag{1}$$

where x is the decision vector, $\Omega \subseteq \mathbb{R}^n$ is the feasible decision space, $f(x)$ is the objective vector defined in the objective space $\mathbb{R}^m$, $c(x) = 0$ are the equality constraints and $g(x) \geq 0$ are the inequality constraints.

When several objectives are optimized at the same time, the search space becomes partially ordered. In such scenario, solutions are compared on the basis of the Pareto dominance. For two solutions a and b from Ω, a solution a is said to dominate a solution b (denoted by $a \prec b$) if:

$$\begin{array}{ll} \forall i \in \{1, \ldots, m\}: & f_i(a) \leq f_i(b) \wedge \\ \exists j \in \{1, \ldots, m\}: & f_j(a) < f_j(b). \end{array} \tag{2}$$

Since solutions are compared against different objectives, there is no longer a single optimal solution but a set of optimal solutions, generally known as the Pareto optimal set. This set contains equally important solutions representing different trade-offs between the given objectives and can be defined as:

$$\mathcal{PS} = \{x \in \Omega \,|\, \nexists y \in \Omega : y \prec x\}. \tag{3}$$

Approximating the Pareto optimal set is the main goal in multi-objective optimization.

Constraint handling in multi-objective optimization is still a very challenging research endeavor. Constraints are present in every real world problem and, so far, there is no superior approach to the handling of constraints. The inclusion of constraints, in particular, equality constraints, further complexifies multi-objective optimization since its inclusion can greatly transform the Pareto Frontier and, thus, make its approximation much more difficult.

In this paper, a Scalarized Augmented Lagrangian Algorithm (SCAL) for constrained multi-objective optimization problems is presented and tested in several constrained problems.

Costa, L., Santo, I. and Oliveira, P.
A Scalarized Augmented Lagrangian Algorithm (SCAL) for Multi-objective Optimization Constrained Problems.
DOI: 10.5220/0006720603350340
In *Proceedings of the 7th International Conference on Operations Research and Enterprise Systems (ICORES 2018)*, pages 335-340
ISBN: 978-989-758-285-1

2 AUGMENTED WEIGHTED TCHEBYCHEFF METHODS

The main goal of the MO optimization is to obtain the set of Pareto-optimal solutions that correspond to different trade-offs between objectives. In this study, a scalarization method is used to obtain an approximation to the Pareto-optimal set (Miettinen, 1999).

The Augmented Weighted Tchebycheff method, proposed by Steuer and Choo (Steuer and Choo, 1983), has the advantage that it can converge to non-extreme final solutions and may also be applicable to nonlinear and nonconvex multi-objective optimization problems (Steuer and Choo, 1983). The Augmented Weighted Tchebycheff method can be formulated as:

$$\min\max[w_i|f_i(x) - z_i^\star|] + \rho \sum_{i=1}^{m} |f_i(x) - z_i| \qquad (4)$$

where w_i are the weighting coefficients for objective i, $z_i^\star$ are the components of a reference point, and ρ is a small positive value (Dchert et al., 2012). The solutions of the optimization problem defined by Eq. 4 are Pareto-optimal solutions. Different combinations allow computing approximations to the Pareto-optimal set using a single objective optimization method. An approximation to the ideal vector can be used as reference point $z^\star = (z_1^\star, \ldots, z_m^\star)^T = (\min f_1, \ldots, \min f_m)^T$.

3 AUGMENTED LAGRANGIAN TECHNIQUE USING THE HOOKE AND JEEVES PATTERN SEARCH METHOD

The Augmented Lagrangian technique herein presented solves a sequence of simple subproblems where the objective function penalizes all or some of the constraint violation. This objective function is an Augmented Lagrangian that depends on a penalty parameter and on the multiplier vectors and is based in the ideas presented in (Bertsekas, 1999; Conn et al., 1991; Lewis and Torczon, 2002):

$$\begin{aligned}\Phi(x;\lambda,\delta,\mu) \;=\;& f(x) + \lambda^T c(x) + \frac{1}{2\mu}\|c(x)\|^2 \qquad (5)\\ &+ \frac{\mu}{2}\left(\left\|\left[\delta + \frac{g(x)}{\mu}\right]_+\right\|^2 - \|\delta\|^2\right)\end{aligned}$$

where μ is a positive penalty parameter, $\lambda = (\lambda_1, \ldots, \lambda_m)^T$, $\delta = (\delta_1, \ldots, \delta_p)^T$ are the Lagrange multiplier vectors associated with the equality and inequality constraints, respectively. Function Φ aims to penalize solutions that violate the equality and inequality constraints, not including the simple bounds $l \le x \le u$. To force feasibility, the inner iterative process must return an approximate solution that satisfies the bound constraints. While using the Hooke and Jeeves version of the pattern search (Hooke and Jeeves, 1961), any computed solution x that does not satisfy the bounds is projected onto the set Ω component by component (for all $i, \ldots, n$) as follows:

$$x_i = \begin{cases} l_i & \text{if} \quad x_i < l_i \\ x_i & \text{if} \quad l_i \le x_i \le u_i \\ u_i & \text{if} \quad x_i > u_i \end{cases} . \qquad (6)$$

The corresponding subproblem is then formulated as:

$$\underset{x\in\Omega}{\text{minimize}}\ \Phi(x;\lambda^j,\delta^j,\mu^j) \qquad (7)$$

where, for each set of fixed λ^j, δ^j and μ^j, the solution of subproblem (7) provides an approximation x^j to the problem formulated in Eq. (4), where the index j is the iteration counter of the outer iterative process. We refer to (Bertsekas, 1999) for details. In practice, common safeguarded schemes maintain the sequence of penalty parameters far away from zero so that solving subproblem (7) is an easy task.

To evaluate the equality and inequality constraint violation, and the complementarity, the following error function is used:

$$E(x,\delta) = \max\left\{\frac{\|c(x)\|_\infty}{1+\|x\|}, \frac{\|[g(x)]_+\|_\infty}{1+\|\delta\|}, \frac{\max_i \delta_i |g_i(x)|}{1+\|\delta\|}\right\}. \qquad (8)$$

The Lagrange multipliers λ^j and δ^j are estimated in this iterative process using the first-order updating formulae

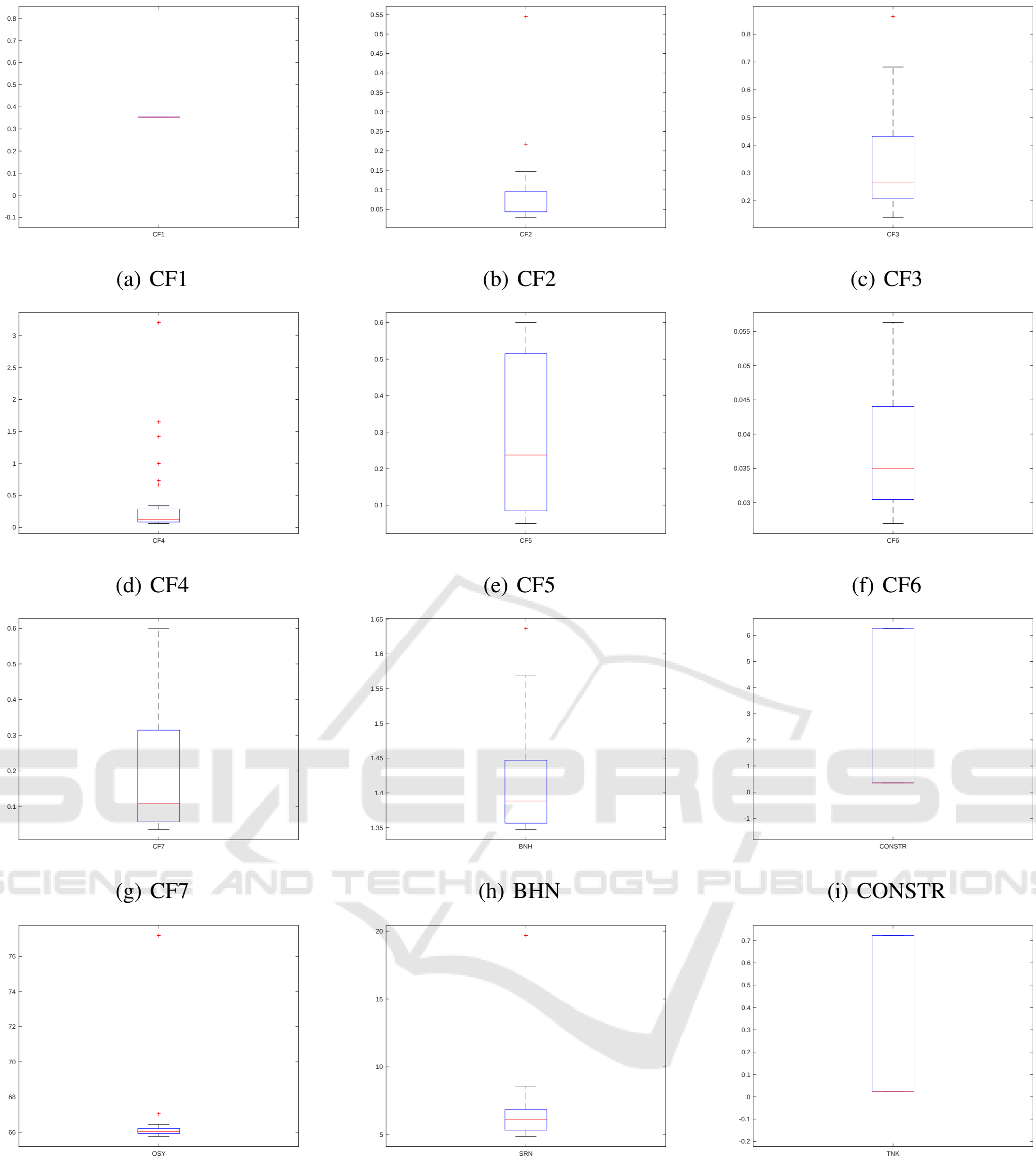
$$\bar{\lambda}_i^{j+1} = \lambda_i^j + \frac{c_i(x^j)}{\mu^j},\ i = 1,\ldots,m \qquad (9)$$

and

$$\bar{\delta}_i^{j+1} = \max\left\{0, \delta_i^j + \frac{g_i(x^j)}{\mu^j}\right\},\ i = 1,\ldots,p \qquad (10)$$

where: for all $j \in \mathbb{N}$, and for $i = 1,\ldots,m$ and $l = 1,\ldots,p$, λ_i^{j+1} is the projection of $\bar{\lambda}_i^{j+1}$ on the interval $[\lambda_{\min}, \lambda_{\max}]$ and δ_i^{j+1} is the projection of $\bar{\delta}_i^{j+1}$ on the interval $[0, \delta_{\max}]$, where $-\infty < \lambda_{\min} \le \lambda_{\max} < \infty$ and $0 \le \delta_{\max} < \infty$. After the new approximation x^j has been computed, the Lagrange multiplier vector δ associated with the inequality constraints is updated, in all iterations, since δ^{j+1} is required in the error function (8) to measure constraint violation and complementarity. We note that the Lagrange multipliers λ_i, $i = 1,\ldots,m$ are updated only when feasibility and complementarity are at a satisfactory level, herein defined by the condition

$$E(x^j, \delta^{j+1}) \le \eta^j \qquad (11)$$



(a) CF1 (b) CF2 (c) CF3

(d) CF4 (e) CF5 (f) CF6

(g) CF7 (h) BHN (i) CONSTR

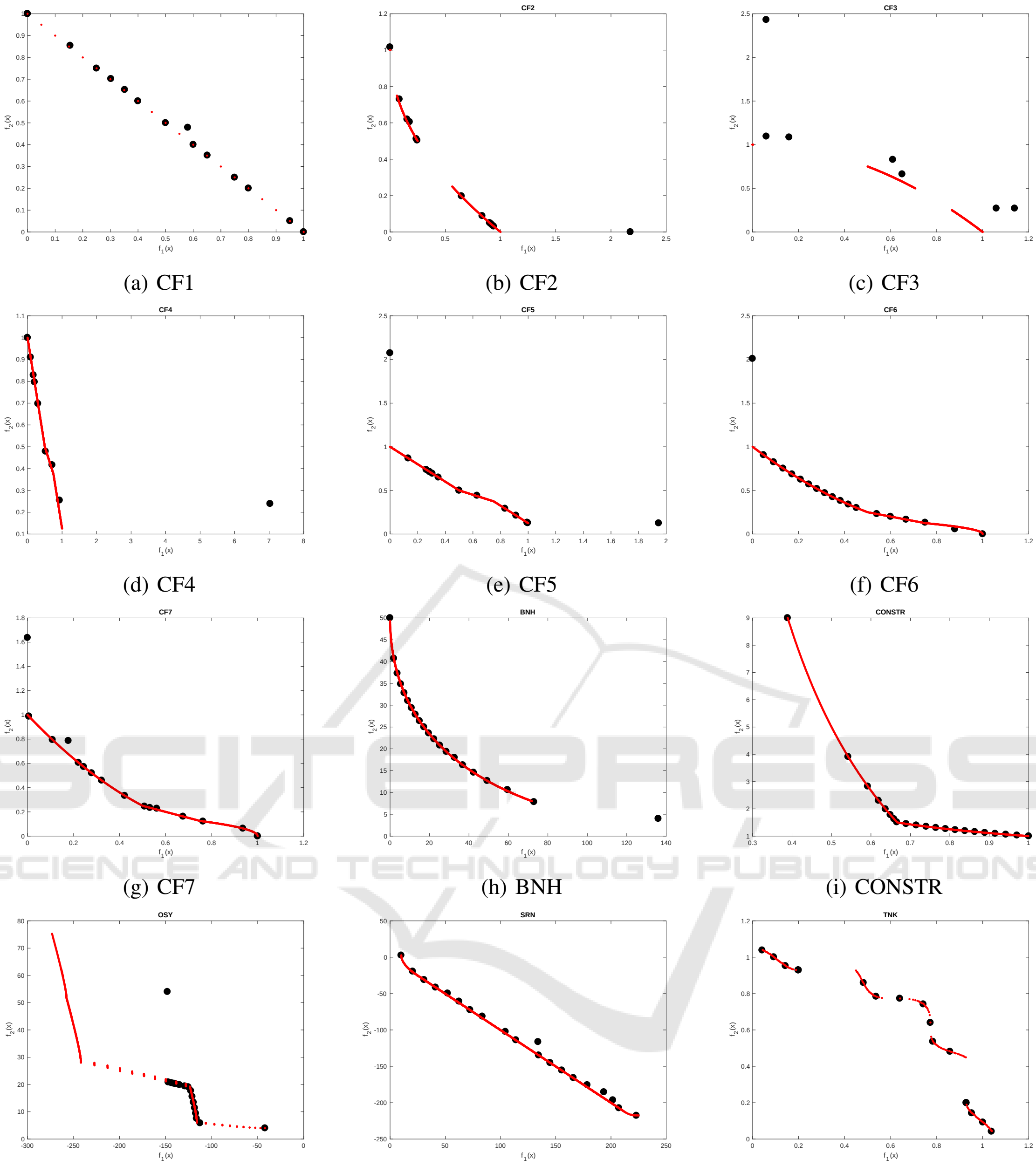
(j) OSY (k) SRN (l) TNK

Figure 1: Boxplots of IGD on test problems.

for a positive tolerance η^j. It is required that $\{\eta^j\}$ defines a decreasing sequence of positive values converging to zero, as $j \to \infty$. This is easily achieved by $\eta^{j+1} = \pi\eta^j$ for $0 < \pi < 1$.

We consider that an iteration j failed to provide an approximation x^j with an appropriate level of feasibility and complementarity if condition (11) does not hold. In this case, the penalty parameter is decreased using $\mu^{j+1} = \gamma\mu^j$ where $0 < \gamma < 1$. When condition (11) holds, then the iteration is considered satisfactory. This condition says that the iterate x^j is feasible and the complementarity condition is satisfied within some tolerance η^j and, consequently, the algorithm maintains the penalty parameter value. We remark that when (11) fails to hold infinitely many times, the sequence of penalty parameters tends to zero. To be able to define an algorithm where the sequence $\{\mu^j\}$ does not reach zero, the following update is used instead:

$$\mu^{j+1} = \max\{\mu_{\min}, \gamma\mu^j\}, \tag{12}$$



(a) CF1 (b) CF2 (c) CF3

(d) CF4 (e) CF5 (f) CF6

(g) CF7 (h) BNH (i) CONSTR

(j) OSY (k) SRN (l) TNK

Figure 2: Illustrations of the Pareto fronts and an example run on test problems.

where $\mu_{\min}$ is a sufficiently small positive real value.

In our algorithm, $\Delta^k s^k$ is computed by the Hooke and Jeeves (HJ) search method (Hooke and Jeeves, 1961). This algorithm differs from the traditional coordinate search since it performs two types of moves: the exploratory move and the pattern move. An exploratory move is a coordinate search – a search along the coordinate axes – around a selected approximation, using a step length Δ^k. A pattern move is a promising direction that is defined by $z^k - z^{k-1}$ when the previous iteration was successful and z^k was accepted as the new approximation. A new trial approximation is then defined as $z^k + (z^k - z^{k-1})$ and an exploratory move is then carried out around this trial point. If this search is successful, the new approximation is accepted as z^{k+1}. We refer to (Hooke and Jeeves, 1961; Lewis and Torczon, 1999) for details. This HJ iterative procedure terminates, providing a new approximation x^j to the problems (4), $x^j \leftarrow z^{k+1}$, when the following stopping condition is satisfied,

Table 1: Performance of MOEA/D-IEpsilon (SBX), MOEA/D-IEpsilon (DE), and SCAL on CF1-CF7, BNH, CONSTR, OSY, SRN and TNK problems in terms of the mean and standard deviation values of IGD.

Problem	MOEA/D-IEpsilon (SBX)	MOEA/D-IEpsilon (DE)	SCAL
CF1 mean	6.06E-03	**5.13E-03**	3.54E-01
std	1.36E-03	9.43E-04	1.13E-16
CF2 mean	1.04E-01	**1.41E-02**	9.49E-02
std	4.62E-02	2.15E-02	9.46E-02
CF3 mean	3.15E-01	**2.33E-01**	3.30E-01
std	1.25E-01	1.47E-01	1.79E-01
CF4 mean	1.18E-01	**2.73E-02**	3.97E-01
std	3.12E-02	9.17E-03	6.65E-01
CF5 mean	2.91E-01	**2.49E-01**	2.94E-01
std	1.33E-01	1.09E-01	2.00E-01
CF6 mean	1.34E-01	6.46E-02	**3.73E-02**
std	7.21E-02	3.59E-02	8.14E-03
CF7 mean	2.48E-01	2.08E-01	**1.96E-01**
std	8.81E-02	8.50E-02	1.77E-01
BNH mean	**3.91E-01**	4.59E-01	1.41E+00
std	1.17E-01	1.11E-01	7.81E-02
CONSTR mean	**8.33E-03**	1.28E-02	2.10E+00
std	9.96E-04	1.29E-03	2.75E+00
OSY mean	**3.86E+00**	4.31E+00	6.51E+00
std	5.60E-01	8.67E-01	2.19E+00
SRN mean	**3.54E-01**	3.56E-01	6.62E+00
std	5.89E-03	8.26E-03	2.63E+00
TNK mean	2.21E-03	**1.82E-03**	3.00E-01
std	9.08E-05	4.47E-05	3.36E-01

$\Delta^k \leq \varepsilon^j$. However, if this condition can not be satisfied in $k_{\max}$ iterations, then the procedure is stopped with the last available approximation.

More detailed explanation on this can be found in (Costa et al., 2012).

4 RESULTS AND DISCUSSION

In order to test the Augmented Lagrangian in constraint handling, several multi-objective test problems are used: CF1-CF7 (Zhang et al., 2009), BNH (Binh and Korn, 1997), CONSTR (Deb, 2001), OSY (Osyczka and Kundu, 1995), SRN (Srinivas and Deb, 1994), and TNK (Tanaka et al., 1995). Furthermore, the results with this new approach are compared with the results of (Fan et al., 2017), in particular, the algorithm MOEA/D-IEpsilon, since this was the best algorithm of all those tested, using simulated binary crossover (SBX) and differential evolution crossover (DE) operators (Fan et al., 2017). The comparison is based on the Inverted Generational Distance (IGD) measure; this measure evaluates algorithm performance in terms of convergence to the Pareto front as well as the diversity of the approximation along the frontier. For this purpose it is mandatory to know the exact definition of the Pareto Frontier. These are preliminary results and no thorough study has been made of the parameters of the Augmented Lagrangian. Figures 1 presents the box plots for 30 executions for each problem and Figure 2 shows the Pareto Frontier (line) for the different problems with the approximation (dots in the graphs) resulting from one of the executions. It can be observed that the algorithm closely approximates the Pareto Frontier, exhibiting some extreme points which have a large impact on the IGD measure. Table 1 presents the results of the new approach and the published results with algorithm MOEA/D-IEpsilon (Fan et al., 2017). Although, the new algorithm only wins in two of the problems, without much investigation of the Augmented Lagrangian parameters, these results seem promising.

5 CONCLUSIONS

In this work a new approach for dealing with constraints in multi-objective optimization problems is proposed based on the Augmented Lagrangian coupled with weighted Tchebycheff method. Although the reported results are a preliminary work on this approach, the results are very encouraging. Therefore, future work will address the study of the parameters of the Augmented Lagrangian, the use of an achievement scalarizing function, and the testing of the algorithm in problems with three objectives.

ACKNOWLEDGEMENTS

This work has been supported by the *Portuguese Foundation for Science and Technology* (FCT) in the scope of the project UID/CEC/00319/2013 (ALGORITMI R&D Center).

REFERENCES

Bertsekas, D. P. (1999). *Nonlinear Programming*. Athena Scientific, Belmont, 2 edition.

Binh, T. T. and Korn, U. (1997). Mobes: A multiobjective evolution strategy for constrained optimization problems. In *In Proceedings of the Third International Conference on Genetic Algorithms*, pages 176–182.

Conn, A. R., Gould, N. I. M., and Toint, P. L. (1991). A globally convergent augmented Lagrangian algorithm for optimization with general constraints and simple bounds. *SIAM Journal on Numerical Analysis*, 28(2):545–572.

Costa, L., Santo, I. A. E., and Fernandes, E. M. (2012). A hybrid genetic pattern search augmented lagrangian method for constrained global optimization. *Applied Mathematics and Computation*, 218(18):9415 – 9426.

Dchert, K., Gorski, J., and Klamroth, K. (2012). An augmented weighted tchebycheff method with adaptively chosen parameters for discrete bicriteria optimization problems. *Computers & Operations Research*, 39(12):2929 – 2943.

Deb, K. (2001). *Multi-Objective Optimization using Evolutionary Algorithms*. Wiley-Interscience Series in Systems and Optimization. John Wiley & Sons.

Fan, Z., Fang, Y., Li, W., Lu, J., Cai, X., and Wei, C. (2017). A comparative study of constrained multi-objective evolutionary algorithms on constrained multi-objective optimization problems. In *2017 IEEE Congress on Evolutionary Computation (CEC)*, pages 209–216.

Hooke, R. and Jeeves, T. A. (1961). Direct search solution of numerical and statistical problems. *Journal on Associated Computation*, 8:212–229.

Lewis, R. and Torczon, V. (1999). Pattern search algorithms for bound constrained minimization. *SIAM Journal on Optimization*, 9(4):1082–1099.

Lewis, R. M. and Torczon, V. (2002). A globally convergent augmented Lagrangian pattern search algorithm for optimization with general constraints and simple bounds. *SIAM Journal on Optimization*, 12(4):1075–1089.

Miettinen, K. (1999). *Nonlinear multiobjective optimization*. Kluwer Academic Publishers, Boston.

Osyczka, A. and Kundu, S. (1995). A new method to solve generalized multicriteria optimization problems using the simple genetic algorithm. *Structural optimization*, 10(2):94–99.

Srinivas, N. and Deb, K. (1994). Muiltiobjective optimization using nondominated sorting in genetic algorithms. *Evolutionary Computation*, 2(3):221–248.

Steuer, R. E. and Choo, E.-U. (1983). An interactive weighted tchebycheff procedure for multiple objective programming. *Mathematical Programming*, 26(3):326–344.

Tanaka, M., Watanabe, H., Furukawa, Y., and Tanino, T. (1995). Ga-based decision support system for multicriteria optimization. In *1995 IEEE International Conference on Systems, Man and Cybernetics. Intelligent Systems for the 21st Century*, volume 2, pages 1556–1561 vol.2.

Zhang, Q., Zhou, A., Zhao, S., Suganthan, P. N., Liu, W., and Tiwari, S. (2009). Multiobjective optimization test instances for the CEC 2009 special session and competition. Technical Report CES-487, University of Essex, UK.