

# Electromagnetism-like Augmented Lagrangian Algorithm for Global Optimization

Ana Maria A. C. Rocha and Edite M. G. P. Fernandes

**Abstract** This paper presents an augmented Lagrangian algorithm to solve continuous constrained global optimization problems. The algorithm approximately solves a sequence of bound constrained subproblems whose objective function penalizes equality and inequality constraints violation and depends on the Lagrange multiplier vectors and a penalty parameter. Each subproblem is solved by a population-based method that uses an electromagnetism-like mechanism to move points towards optimality. Benchmark problems are solved in a performance evaluation of the proposed augmented Lagrangian methodology. A comparison with a well-known technique is also reported.

## 1 Introduction

This paper presents a numerical study of an augmented Lagrangian methodology, where the subproblems are solved by a stochastic population based algorithm, for continuous constrained global optimization. We aim to address the problem in the form:

$$\min f(x) \text{ subject to } g(x) \leq 0, h(x) = 0, x \in \Omega, \quad (1)$$

where  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$  and  $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$  are nonlinear continuous functions and  $\Omega = \{x \in \mathbb{R}^n : lb \leq x \leq ub\}$ . We do not assume that the objective function  $f$  is convex. There may be many local minima in the feasible region. This class of global optimization problems arises frequently in engineering applications. Spe-

---

Ana Maria A. C. Rocha

Department of Production and Systems, University of Minho, 4710-057 Braga, Portugal e-mail: arocha@dps.uminho.pt

Edite M. G. P. Fernandes

Department of Production and Systems, University of Minho, 4710-057 Braga, Portugal e-mail: emgpf@dps.uminho.pt

cially for large scale problems, derivative-free and stochastic methods are the most well-known and used methods. The two main categories of methods to handle constraints in these algorithms are listed below.

1. Methods based on penalty functions. The constraints violation is combined with the objective function to define a penalty function. This function aims to penalize infeasible solutions by increasing their fitness values proportionally to their level of constraints violation. Static, dynamic, annealing and adaptive penalties are the most popular [2, 4, 8, 14, 16, 19]. Methods based on augmented Lagrangians are common in deterministic type methods for global optimization, for example in [6, 7, 15], but rare when combined with heuristics that rely on a population of points to converge to the solution [1, 22, 23, 24].
2. Methods based on biasing feasible over infeasible solutions. They seem to be nowadays interesting alternatives to penalty methods for handling constraints. In this type of methods, constraints violation and the objective function are used separately and optimized by some sort of order, being the constraints violation the most important. See, for example, [9, 17, 18, 20, 21].

Here, we aim to show the functionality of an augmented Lagrangian methodology to handle the equality and inequality constraints of the problem (1), where the subproblems are approximately solved by a stochastic global population based algorithm. Due to its simplicity, the electromagnetism-like (EM) algorithm proposed in [4, 5] is used to obtain the solution of each subproblem. The EM algorithm simulates the electromagnetism theory of physics by considering each point in the population as an electrical charge. The method uses an attraction-repulsion mechanism to move a population of points towards optimality.

Since the EM algorithm has been designed to find a minimizer which satisfies  $x \in \Omega$ , our subproblem has bound constraints. Although other constraint-handling techniques have been implemented by the authors with the EM algorithm, namely, the feasibility and dominance rules [17, 18], the separate feasibility and optimality measures based on sufficient reduction conditions [20], and the adaptive penalty technique [19], the new proposed augmented Lagrangian strategy has been given the best results so far.

Our implementation of an augmented Lagrangian methodology is as follows: we reformulate problem (1) converting each equality constraint into an inequality as herein shown:  $|h_j(x)| \leq \beta$ , where  $\beta$  is a positive relaxation parameter. This is an usual procedure in stochastic based methods [9, 18, 19]. In general, the relaxation parameter is fixed over the entire iterative process. Typically,  $10^{-3}$ ,  $10^{-4}$  and  $10^{-5}$  are common values in the literature. Our proposal defines a sequence  $\{\beta^k\}$  of decreasing nonnegative numbers such that  $\lim_{k \rightarrow \infty} \beta^k = \beta^* > 0$ . The idea is to tighten the equality constraints relaxation scheme as iterations proceed. Further, a different updating scheme for the penalty parameter is also proposed. When the level of constraints violation is under a specified tolerance, even if the infeasibility did not improve, the penalty is allowed to decrease instead of increasing (see Algorithm 2.1).

The remainder of this paper is organized as follows. Sect. 2 describes the proposed stochastic augmented Lagrangian paradigm and Sect. 3 briefly introduces the

EM algorithm. Sect. 4 displays the results of the numerical experiments, including a comparison with a well-known penalty method, and Sect. 5 contains some concluding remarks.

## 2 An augmented Lagrangian method

Most stochastic methods for global optimization are developed primarily for unconstrained or simple bound constrained problems. Then they are extended to constrained optimization problems using, for example, a penalty technique. This type of technique transforms the constrained problem into a sequence of unconstrained subproblems by penalizing the objective function when constraints are violated. The objective penalty function, in the unconstrained subproblem, depends on a positive penalty parameter that must be updated throughout the iterative process. With most penalty functions, the solution of the constrained problem is reached for an infinite value of the penalty parameter. An augmented Lagrangian is a more sophisticated penalty function for which a finite penalty parameter value is sufficient to yield convergence to the solution of the constrained problem [3].

We now show the functionality of the proposed augmented Lagrangian function when solving constrained global optimization problems. Practical and theoretical issues from the augmented Lagrangian methodology are used with this population based algorithm, the EM algorithm as proposed in [5], to compute approximate solutions of the sequence of bound constrained subproblems.

Since equality constraints are the most difficult to be satisfied, our augmented Lagrangian methodology considers problems only with inequality constraints, using a common procedure in stochastic methods for global optimization to convert the equality constraints of the problem into inequality constraints, as follows:  $|h_j| \leq \beta$ ,  $j = 1, \dots, m$  for a fixed  $\beta > 0$ . For simplicity, the problem (1) is rewritten as

$$\min f(x) \text{ subject to } G(x) \leq 0, x \in \Omega, \quad (2)$$

where the vector of the inequality constraints is now defined by

$$G(x) = (g_1(x), \dots, g_p(x), |h_1(x)| - \beta, \dots, |h_m(x)| - \beta).$$

Our proposal concerning the relaxed equality constraints aims to tighten the relaxation scheme as iterations proceed, using variable relaxation parameter values. Thus, a sequence of decreasing nonnegative values bounded by  $\beta^* > 0$  is defined as:

$$\beta^{k+1} = \max \left\{ \beta^*, \frac{1}{\sigma} \beta^k \right\}, \sigma > 1. \quad (3)$$

The formula of the augmented Lagrangian that corresponds to the inequality constraints in the converted problem (2) is:

$$\mathcal{L}_\rho(x, \mu) = f(x) + \frac{\rho}{2} \sum_{i=1}^{p+m} \left[ \max \left( 0, G_i(x) + \frac{\mu_i}{\rho} \right) \right]^2 \quad (4)$$

where  $\rho$  is a positive penalty parameter and  $\mu \in \mathbb{R}^{p+m}$  represents the Lagrange multiplier vector associated with the  $p + m$  constraints. Our proposed stochastic augmented Lagrangian algorithm adapted to the reformulation (2), of the original problem (1), and based on the Lagrangian (4) is presented in Algorithm 2.1:

---

**Algorithm 2.1** (Augmented Lagrangian algorithm)

---

```

1: Given:  $\mu^+ > 0, \varepsilon^* > 0, 0 < \alpha < 1, \gamma > 1, \sigma > 1, k_{\max}, l_{\max}, \beta^* > 0, 0 < \rho^- < \rho^+, \mu^1 \in [0, \mu^+]$ 
2: randomly generate  $x^0$  in  $\Omega$ ; compute  $\rho^1$ ; set  $k = 1$ 
3: while  $\|v^{k-1}\| > \varepsilon^*$  and  $k \leq k_{\max}$  do
4:    $\varepsilon^k = \max\{\varepsilon^*, 10^{-k}\}$ ; update  $\beta^k$  using (3); set  $l = 1$ 
5:   while  $(\mathcal{L}_{\text{avg}} - \mathcal{L}_{\rho^k}(x(\text{best}), \mu^k)) > \varepsilon^k$  and  $l \leq l_{\max}$  do
6:     use  $x^{k-1}$  and randomly initialize a population of  $p_{\text{size}} - 1$  points in  $\Omega$ 
7:     run EM to compute a population of solutions to  $\min_x \mathcal{L}_{\rho^k}^l(x, \mu^k)$  subject to  $x \in \Omega$ 
8:      $l = l + 1$ 
9:   end while
10:   $x^k = x(\text{best})$ 
11:  compute  $v_i^k = \max\left\{G_i(x^k), -\frac{\mu_i^k}{\rho^k}\right\}, i = 1, \dots, m + p$ 
12:  if  $k = 1$  or  $\|v^k\| \leq \alpha \|v^{k-1}\|$  then
13:     $\rho^{k+1} = \rho^k$ 
14:  else
15:    if  $\|v^k\| \leq \varepsilon^k$  then
16:       $\rho^{k+1} = \max\{\rho^-, \frac{1}{\gamma} \rho^k\}$ 
17:    else
18:       $\rho^{k+1} = \min\{\rho^+, \gamma \rho^k\}$ 
19:    end if
20:  end if
21:  update  $\mu_i^{k+1} = \min\{\max\{0, \mu_i^k + \rho^k G_i(x^k)\}, \mu^+\}, i = 1, \dots, m + p$ 
22:   $k = k + 1$ 
23: end while

```

---

This algorithm extends recent work with the Powell-Hestenes-Rockafellar augmented Lagrangian function presented in [6, 7], where equalities and inequalities are treated separately. In Algorithm 2.1 the penalty parameter  $\rho$ , besides being increased, when infeasibility is not reduced, it is also reduced whenever the constraints violation is under a specified tolerance  $\varepsilon^k$ ; otherwise it is not changed. Further, a safeguarded scheme is also included in the process. This is motivated by the need to keep the penalty parameter bounded and the subproblems well conditioned. This procedure is reported in the lines 12–20 of Algorithm 2.1. The initial value for the penalty parameter is defined by

$$\rho^1 = \max \left\{ 10^{-6}, \min \left\{ 10, 2|f(x^0)| / (\|\max(0, G(x^0))\|^2) \right\} \right\}$$

for an arbitrary initial approximation  $x^0$  [6, 7]. The algorithm also updates the Lagrange multipliers using first order estimates and safeguarded schemes. This is a crucial issue to maintain the sequence  $\{\mu^k\}$  bounded.

Further, lines 5-9 of the algorithm show details of the inner iterative process to compute an approximation to the solution of subproblem (5), at each iteration  $k$  :

$$\min_x \mathcal{L}_{\rho^k}(x, \mu^k) \text{ subject to } x \in \Omega \quad (5)$$

for fixed values of the parameters  $\rho^k$  and  $\mu^k$ . For each set of fixed values of penalties and Lagrange multipliers, a reasonable approximate solution of subproblem (5) is required so that convergence could be promoted [3]. Since the EM algorithm is based on a population of points, with size  $p_{\text{size}}$ , the point which yields the least objective function value, denoted by the best point of the population,  $x(\text{best})$ , after stopping, is taken as the next approximation to the problem (1). We also note that our stochastic EM algorithm uses the approximation  $x^{k-1}$  as one of the points of the population to initialize the EM algorithm. The remaining  $p_{\text{size}} - 1$  points are randomly generated. The inner iteration counter is represented by  $l$ . This process terminates when the difference between the function value at the best point,  $\mathcal{L}_{\rho^k}(x(\text{best}), \mu^k)$ , and the average of the function values of the population,  $\mathcal{L}_{\text{avg}}$ , is under a specified tolerance  $\varepsilon^k$ . This tolerance decreases as outer iterations proceed. A limit of  $l_{\text{max}}$  iterations is also imposed.

### 3 The electromagnetism-like mechanism

In this section, we briefly present the EM mechanism, proposed in [5], for solving the subproblems in the Algorithm 2.1. Here, the objective is to compute an approximate minimizer of  $\mathcal{L}_{\rho^k}(x, \mu^k)$ , for fixed values of the parameters  $\rho^k$  and  $\mu^k$ . For simplicity we use the notation  $\mathcal{L}^k(x) = \mathcal{L}_{\rho^k}(x, \mu^k)$ . Because EM is a population based algorithm, the inner iterative process begins with a population of  $p_{\text{size}}$  solutions (line 6 in Algorithm 2.1). The best found solution, denoted by  $x(\text{best})$ , and the average of function values, are defined by

$$x(\text{best}) = \arg \min \left\{ \mathcal{L}^k(x(s)) : s = 1, \dots, p_{\text{size}} \right\} \text{ and } \mathcal{L}_{\text{avg}}^k = \sum_{s=1}^{p_{\text{size}}} \mathcal{L}^k(x(s)) / p_{\text{size}}, \quad (6)$$

respectively, where  $x(s), s = 1, \dots, p_{\text{size}}$  represent the points of the population. The main steps of the EM mechanism are shown in Algorithm 3.1. Details of each step follow.

The EM mechanism starts by identifying the best point,  $x(\text{best})$ , of the population using the augmented Lagrangian  $\mathcal{L}^k$  for point assessment, see (6). According to the electromagnetism theory, the total force exerted on each point  $x(s)$  by the other  $p_{\text{size}} - 1$  points is inversely proportional to the square of the distance between the points and directly proportional to the product of their charges:

**Algorithm 3.1** (EM algorithm)

- 
- 1: **Given:**  $x(s), s = 1, \dots, p_{\text{size}}$
  - 2: evaluate the population and select  $x(\text{best})$
  - 3: compute the charges  $c(s), s = 1, \dots, p_{\text{size}}$
  - 4: compute the total forces  $F(s), s = 1, \dots, p_{\text{size}}$
  - 5: move the points except  $x(\text{best})$
  - 6: evaluate the new population and select  $x(\text{best})$
  - 7: apply local search to  $x(\text{best})$
  - 8: compute  $\mathcal{L}^k(x(\text{best}))$  and  $\mathcal{L}_{\text{avg}}^k$ .
- 

$$F(s) = \sum_{r \neq s}^{p_{\text{size}}} F_r^s \equiv \begin{cases} (x(r) - x(s)) \frac{c(s)c(r)}{\|x(r) - x(s)\|^2}, & \text{if } \mathcal{L}^k(x(r)) < \mathcal{L}^k(x(s)) \\ (x(s) - x(r)) \frac{c(s)c(r)}{\|x(r) - x(s)\|^2}, & \text{otherwise} \end{cases},$$

for  $s = 1, \dots, p_{\text{size}}$ , where the charge  $c(s)$  of point  $x(s)$  determines the magnitude of attraction of that point over the others through

$$c(s) = \exp \left( \frac{-n (\mathcal{L}^k(x(s)) - \mathcal{L}^k(x(\text{best})))}{\sum_{r=1}^{p_{\text{size}}} (\mathcal{L}^k(x(r)) - \mathcal{L}^k(x(\text{best})))} \right).$$

Then, the normalized total force vector exerted on each point  $x(s)$  is used to move the point in the direction of the force by a random step size  $\iota \sim U[0, 1]$ , maintaining the point inside the set  $\Omega$ . Thus for  $s = 1, \dots, p_{\text{size}}$  ( $s \neq \text{best}$ ) and for each component  $i = 1, \dots, n$

$$x_i(s) = \begin{cases} x_i(s) + \iota \frac{F_i(s)}{\|F(s)\|} (ub_i - x_i(s)), & \text{if } F_i(s) > 0 \\ x_i(s) + \iota \frac{F_i(s)}{\|F(s)\|} (x_i(s) - lb_i), & \text{otherwise} \end{cases}.$$

Finally, a local search is performed around the best point of the population in order to refine the solution. In [5], a simple random local search is proposed. This is described in Subsect. 3.1. In this paper, we also aim to integrate in the EM algorithm a more sophisticated local search (see Subsect. 3.2). Our numerical experiments show that our proposal significantly improves the performance of the overall augmented Lagrangian algorithm.

### 3.1 Random local search

Here, we briefly describe the local search proposed in [5] for the population-based EM algorithm. This is a simple random line search applied component by component to  $x(\text{best})$ . For each component  $i$ ,  $x(\text{best})$  is assigned to a temporary point  $y$ .

Then a random movement of maximum length  $\delta_{EM} \max_j(ub_j - lb_j)$ ,  $\delta_{EM} > 0$  is carried out and if a better position is obtained within  $\max_{local}$  iterations,  $x(\text{best})$  is replaced by  $y$ , the search ends for that component and proceeds to another one. Although this local search is based on a simple random procedure, it has been shown to greatly improve accuracy of the EM algorithm.

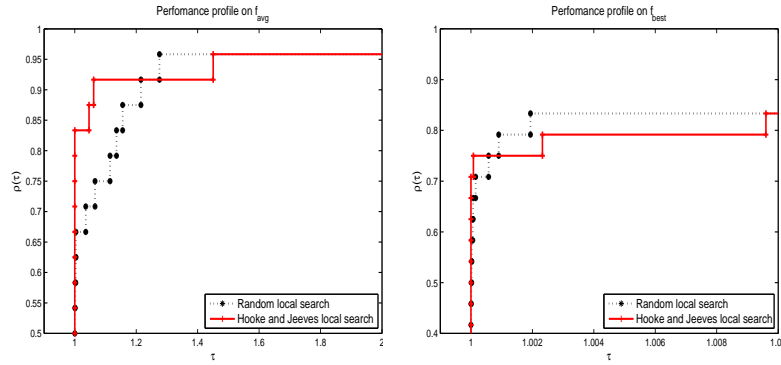
### 3.2 Hooke and Jeeves local search

In this section, we describe our modification to the original EM algorithm. In our proposal, the local procedure is based on the Hooke and Jeeves (HJ) pattern search algorithm. This is a derivative-free method that searches in the neighborhood of a point  $x$  for a better approximation *via* exploratory and pattern moves [11, 12]. To reduce the number of function evaluations, the HJ pattern search algorithm is applied to the current best point only,  $x(\text{best})$ . This algorithm is a variant of the coordinate search, in the sense that incorporates a pattern move to accelerate the progress of the algorithm, by exploiting information obtained from the search in previous successful iterations.

The exploratory move carries out a coordinate search (a search along the coordinate axis) about the best point, with a step length  $\delta_{HJ}$ . If a new trial point,  $y$ , with a better fitness value than  $x(\text{best})$  is encountered, in the sense that the augmented Lagrangian  $\mathcal{L}^k$  value is better, the iteration is successful. Otherwise, the iteration is unsuccessful and  $\delta_{HJ}$  is reduced by a factor  $0 < \Delta_{HJ} < 1$ . If the previous iteration was successful, the vector  $y - x(\text{best})$  defines a promising direction and a pattern move is then implemented, meaning that the exploratory move is carried out about the trial point  $y + (y - x(\text{best}))$ , rather than about the current point  $y$ . Then, if the coordinate search is successful, the returned point is accepted as the new point; otherwise, the pattern move is rejected and the method reduces to coordinate search about  $y$ . Please see [11] for details. To ensure that only points inside  $\Omega$  are tested in the HJ pattern search algorithm, a projection scheme is used, i.e., for each component  $j$ , if  $y_j < lb_j$  or  $y_j > ub_j$  then  $y_j$  is set to  $lb_j$  or  $ub_j$  respectively.

## 4 Numerical Experiments

In this section, we report the results of our numerical study, after running a set of 24 benchmark constrained global problems, described in full detail in [13]. The problems are known as g01-g24 (the ‘g’ suit, where six problems only have equality constraints, thirteen have inequality constraints, five have both equalities and inequalities and all have simple bounds). Not all problems have multi-modal objective functions, although some are difficult to solve. The best known solution for problem g20 is slightly infeasible. We remark that g02, g03, g08 and g12 are maximization problems that were transformed and solved as minimization ones.



**Fig. 1** Comparison of Hooke and Jeeves and Random searches based on  $f_{avg}$  (left) and  $f_{best}$  (right).

Since the algorithm relies on some random parameters and variables, we solve each problem 30 times and take average of the obtained solutions, herein denoted by  $f_{avg}$ . The best of the solutions found after all runs is denoted by  $f_{best}$ . The size of the population depends on  $n$ , and since some problems have large dimension,  $n > 20$ , we choose  $p_{size} = \min\{200, 10n\}$ . The fixed parameters are set in this study as follows:  $\mu^+ = \rho^+ = 10^{12}$ ,  $\epsilon^* = 10^{-6}$ ,  $\alpha = 0.5$ ,  $\gamma = 2$ ,  $\sigma = 2$ ,  $\beta^* = \rho^- = 10^{-12}$ ,  $\beta^1 = 10^{-3}$ . We define  $k_{max} = 50$  and  $l_{max} = 30$  so that a maximum of 1500 iterations are allowed. We remark that the other conditions in the stopping criteria of the Algorithm 2.1 (in the outer and inner iterative processes) may cause the termination of the algorithm before reaching the 1500 iterations. The initial multiplier vector is set to the null vector. The values for the two parameters of the random local search are set as proposed in [5]:  $\max_{local} = 10$ ,  $\delta_{EM} = 0.001$ . In our implementation of the HJ pattern search, we set the initial  $\delta_{HJ}$  to 1 and the factor  $\Delta_{HJ} = 0.1$ .

To compare the performance of the HJ local search procedure, with the random local search of the EM algorithm [5], when incorporated in the proposed augmented Lagrangian context, we use Dolan and Moré's performance profiles [10]. The two profiles are based on the metrics:  $f_{avg}$ , the average of the solutions obtained after the 30 runs, and  $f_{best}$ , the best solution over the 30 runs. For each solver in comparison, the plot shows the proportion of problems in the set, denoted by  $\rho(\tau)$ , that has the best value of the metric, for each value of  $\tau \in \mathbb{R}$ . Thus, to see which solver has the least value of the metric mostly, then its value of  $\rho(1)$  should be compared with those of the other solvers in comparison. The higher the  $\rho$  the better the solver is. On the other hand,  $\rho_s(\tau)$  for large values of  $\tau$  measures the solver robustness.

Hooke and Jeeves local search performs significantly better than Random local search when comparing the average performance (left plot in Fig. 1) and it gives the best results for about 71% of the problems against 57% of the problems with Random search (plot on the right of Fig. 1).

Table 1 reports  $f_{best}$  and  $f_{avg}$  obtained by our study and those of [2] for the eleven problems therein registered (g01-g11) and  $f^*$  is the best known solution as reported



in [13]. In [2], a genetic algorithm combined with an adaptive penalty function is implemented. These are mainly concerned with the frequency of penalty parameters updating and constraints violation computation. For a fair comparison, we use the same conditions described in [2]:  $p_{\text{size}} = 100$ , runs = 25, maximum number of generations = 1000, leading to 100000 fitness function evaluations. We have better performance (both in  $f_{\text{best}}$  and  $f_{\text{avg}}$ ) than the adaptive penalty algorithm of [2] (best results are in boldface) in all but g02, g03, g06 and g10 problems.

**Table 1** Comparison of our results with the best of 5 variants in [2]

Prob.	$f^*$	our study		[2]	
		$f_{\text{best}}$	$f_{\text{avg}}$	$f_{\text{best}}$	$f_{\text{avg}}$
g01	-15.0000	<b>-15.0000</b>	<b>-15.0000</b>	-14.9998	-14.9989
g02	-0.80362	-0.76250	-0.63387	<b>-0.79252</b>	<b>-0.72555</b>
g03	-1.00050	-0.99634	<b>-0.92664</b>	<b>-0.99725</b>	-0.77797
g04	-30665.54	<b>-30665.55</b>	<b>-30665.55</b>	-30665.32	-30578.55
g05	5126.497	<b>5126.497</b>	<b>5126.498</b>	5126.779	5323.866
g06	-6961.814	-6945.097	-6761.717	<b>-6961.448</b>	<b>-6805.229</b>
g07	24.3062	<b>24.3062</b>	<b>24.3062</b>	24.5450	27.8486
g08	-0.09583	<b>-0.09583</b>	<b>-0.08992</b>	<b>-0.09583</b>	-0.08769
g09	680.630	<b>680.630</b>	<b>680.691</b>	680.681	681.470
g10	7049.25	7074.93	<b>7217.08</b>	<b>7070.56</b>	8063.29
g11	0.74990	<b>0.75000</b>	<b>0.75000</b>	0.75217	0.88793

## 5 Final remarks

From our preliminary numerical tests, we may conclude that the proposed augmented Lagrangian algorithm is able to effectively solve constrained problems till optimality and seems to be competitive with a well-known penalty based algorithm. Practical engineering problems, for example, those reported in [2], will be solved in the near future. We also aim to test our algorithm with a point-to-point search yet stochastic method, when solving the bound constrained subproblems.

## References

1. H.J.C. Barbosa. (1999) A coevolutionary genetic algorithm for constrained optimization, in *Proceedings of the 1999 Congress on Evolutionary Computation*, DOI:10.1109/CEC.1999.785466, 3:1605–611.
2. H.J.C. Barbosa and A.C.C. Lemonge. (2008) An adaptive penalty method for genetic algorithms in constrained optimization problems, in *Frontiers in Evolutionary Robotics*, H. Iba (ed.) 34 pages. 2008 (ISBN: 978-3-902613-19-6) I-Tech Education Publ., Austria.
3. D.P. Bertsekas. (1999) *Nonlinear Programming*, 2nd edn. Athena Scientific, Belmont.

4. S.I. Birbil. (2002) *Stochastic Global Optimization Techniques*, Ph.D. diss., North Carolina State University.
5. S.I. Birbil and S.-C. Fang. (2003) An electromagnetism-like mechanism for global optimization, *Journal of Global Optimization* 25:263–282.
6. E.G. Birgin, R. Castillo, and J.M. Martinez. (2004) Numerical comparison of Augmented Lagrangian algorithms for nonconvex problems. *Computational Optimization and Applications* 31:31–56.
7. E.G. Birgin, C.A. Floudas and J.M. Martinez. (2010) Global minimization using an Augmented Lagrangian method with variable lower-level constraints. *Mathematical Programming A* 125(1):139–162.
8. C.A. Coello Coello. (2000) Use of a self-adaptive penalty approach for engineering optimization problems. *Computers in Industry* 41:113–127.
9. K. Deb. (1998) An efficient constraint handling method for genetic algorithms, *Computer Methods in Applied Mechanics and Engineering* 186:311–338.
10. E.D. Dolan and J.J. Moré. (2001) Benchmarking optimization software with performance profiles. *Mathematical Programming* 91:201–213.
11. R. Hooke and T.A. Jeeves. (1961) Direct search solution of numerical and statistical problems. *Journal of Association and Computing Machinery* 8:212–229.
12. R.M. Lewis and V. Torczon. (1999) Pattern search algorithms for bound constrained minimization. *SIAM Journal on Optimization* 9:1082–1099.
13. J.J. Liang, T.P. Runarsson, E. Mezura-Montes, M. Clerc, P.N. Suganthan, C.A.C. Coello, and K. Deb. (2005) *Problem Definitions and Evaluation Criteria for the CEC 2006 Special Session on Constrained Real-Parameter Optimization*, Technical Report, Nanyang Technological University, Singapore.
14. J.-L. Liu and J.-H. Lin. (2007) Evolutionary computation of unconstrained and constrained problems using a novel momentum-type particle swarm optimization. *Engineering Optimization* 39:287–305.
15. H. Luo, X. Sun, and H. Wu. (2008) Convergence properties of augmented Lagrangian methods for constrained global optimization. *Optimization Methods and Software* 23:763–778.
16. Y.G. Petalas, K.E. Parsopoulos, and M.N. Vrahatis. (2007) Memetic particle swarm optimization. *Annals of Operations Research* 156:99–127.
17. A.M.A.C. Rocha and E.M.G.P. Fernandes. (2008) Feasibility and dominance rules in the electromagnetism-like algorithm for constrained global optimization problems. *Lecture Notes in Computer Science Computational Science and Its Applications* (O. Gervasi et al. (eds.)) 5073:768–783.
18. A.M.A.C. Rocha and E.M.G.P. Fernandes. (2008) Implementation of the electromagnetism-like algorithm with a constraint-handling technique for engineering optimization problems, in *2008 Eighth International Conference on Hybrid Intelligent Systems* ISBN: 978-0-7695-3326-1, IEEE Computer Society 690–695.
19. A.M.A.C. Rocha and E.M.G.P. Fernandes. (2009) Self-adaptive penalties in the electromagnetism-like algorithm for constrained global optimization problems, in *Proceedings of the 8th World Congress on Structural and Multidisciplinary Optimization*, CD 10 pages, Lisbon.
20. A.M.A.C. Rocha and E.M.G.P. Fernandes. (2009) Hybridizing the electromagnetism-like algorithm with descent search for solving engineering design problems. *International Journal of Computer Mathematics* 86:1932–1946.
21. T.P. Runarsson and X. Yao. (2000) Stochastic ranking for constrained evolutionary optimization. *IEEE Transactions on Evolutionary Computation* 4:284–294.
22. K. Sedlaczek and P. Eberhard. (2007) Augmented Lagrangian particle swarm optimization in mechanism design. *Journal of System Design and Dynamics* 1:410–421.
23. B.W. Wah and T. Wang. (1999) Efficient and Adaptive Lagrange-Multipliers. *Journal of Global Optimization* 14:1–25.
24. T. Wang and B.W. Wah. (1996) Handling inequality constraints in continuous nonlinear global optimization, in *Integrated Design and Process Science* 267–274.