UNIVERSITY
OF TRENTO

DIPARTIMENTO DI INGEGNERIA E SCIENZA DELL'INFORMAZIONE

38123 Povo – Trento (Italy), Via Sommarive 14
http://www.disi.unitn.it

EVOLUTIONARY OPTIMIZATION AS APPLIED TO INVERSE SCATTERING PROBLEMS

P. Rocca, M. Benedetti, M. Donelli, D. Franceschini, and A. Massa

December 2009

Technical Report # DISI-11-019

Evolutionary Optimization as Applied to Inverse Scattering Problems

P. Rocca, M. Benedetti, M. Donelli, D. Franceschini, and A. Massa

Department of Information Engineering and Computer Science, *ELEDIA* Research Group, University of Trento, Via Sommarive 14, 38050 Trento - Italy, Tel. +39 0461 882057, Fax. +39 0461 882093

E-mail: *andrea.massa@ing.unitn.it*

Abstract. This paper is aimed at presenting an overview of Evolutionary Algorithms (*EA*s) as applied to the solution of inverse scattering problems. The focus of this work is on the use of different population-based optimization algorithms for the reconstruction of unknown objects embedded in an inaccessible region when illuminated by a set of microwaves. Starting from a general description of the structure of *EA*s, the classical stochastic operators responsible for the evolution process are described. The extension to hybrid implementations when integrated with local search techniques and the exploitation of the "domain knowledge", either *a-priori* obtained or collected during the optimization process, are also presented. Some theoretical discussions concerned with the convergence issues and a sensitivity analysis on the parameters influencing the stochastic process are reported, as well. Successively, a review on how various researchers have applied or customized different evolutionary approaches to inverse scattering problems is carried out ranging from the shape reconstruction of perfectly conducting objects to the detection of the dielectric properties of unknown scatterers up to applications to sub-surface or biomedical imaging. Finally, open problems and envisaged developments are discussed.

Key Words **-** Evolutionary algorithms, Inverse scattering.

Classification Numbers (MSC) - 45Q05, 78A46, 78M50, 78M99

1. Introduction

Optimization techniques are generally classified into deterministic and stochastic methods. For example, the greedy, the steepest descent, and the tree search algorithms

[111][119] belong to the former class. Although effective in terms of convergence speed, these methods generally require a "domain knowledge" since in case of non-linear and multi-mimina functionals the initial trial solution must lie in the so-called "attraction basin" of the global solution to avoid the convergence solution being trapped into local minima of the functional (i.e., wrong solutions of the problem at hand). On the contrary, stochastic algorithms [59][144][136] are global search approaches potentially able to find the global optimum of the functional whatever the initial point/s of the search.

The goal of optimization is the knowledge of the global solution. The solution is fully described when its descriptors (i.e., its descriptive features), which quantify the information content of the solution itself, are defined. This can be mathematically done by determining the problem unknowns (i.e., the coded representation of the solution descriptors) through the optimization of a suitable cost function. It should be observed that, the number of unknowns is different in each problem and proportional to the information content of the solution.

Since on one hand the descriptors are different (e.g., discrete/continuous variables) as well as the number of unknowns to be determined can vary among the optimization problems, the choice of a proper optimization algorithm is a key issue and a general rule for this choice does not exist. From a practical point of view, the main features necessary to an optimization algorithm are the ability to deal with complex functionals or cost functions, the "simplicity of use", a limited number of control parameters, good convergence properties, and the exploitation of the parallelism offered by modern PC clusters. In this sense, Evolutionary Algorithms (EAs) seem to be good candidates. They have been applied to a huge variety of problems in different and very heterogeneous fields ranging from engineering to economics, up to business and natural science.

For example, in biomedical and natural science, several researches are concerned with the use of evolutionary algorithms for the prediction of protein structures [84] and the design of drugs [91]. In the framework of engineering, they have been applied to the design of aircrafts [19], the synthesis of electromagnetic systems [161] and microwave devices [160]. As far as combinatorial optimization is concerned, routing [135], assignment [45], and scheduling [102] problems have been dealt with, as well. Although the number of works published in economy and business is limited, evolutionary algorithms have demonstrated to work effectively as shown in [74] and [70].

This work is aimed at discussing the use of evolutionary algorithms on a class of problems in electromagnetic engineering, namely the inverse scattering problems.

The first population based algorithm applied to this topic was the Genetic Algorithm (GA) [39]. Several versions of GAs have been implemented and effectively used in electromagnetic inversion [100][24][174][79]. In order to cope with the drawbacks of GAs, different kinds of evolutionary algorithms has been successively developed. Among them, let us recall the differential evolution (DE) algorithm [130][97] and the particle swarm optimizer (PSO) [49][81]. More recently, the ant colony optimizer (ACO) has been also applied [118]. Besides "bare" techniques, a non-negligible number of hybrid approaches has been implemented to improve the convergence rate of global optimizers.

The paper is organized as follows. In Section 2, an introduction on the genesis of nature-inspired optimization algorithms and some motivations on their use and efficiency when dealing with real world problems are given. A general description of the structure of EAs is presented in Sect. 3, while different implementations are detailed in Sect. 4. Section 5 is devoted to some theoretical discussions on the convergence properties as well as on EAs sensitivity to the values of the control parameters. The inverse scattering problems is briefly formulated in Sect. 6 and an overview on the application of EAs is provided. Some conclusions on the role of EAs in inverse scattering are drawn in Sect. 7, whereas open problems and possible future developments are discussed in Sect. 8.

2. The Origin of EAs - Adaptation in Artificial Systems

In early 1970s [75], Holland showed that nature-inspired evolutionary algorithms can be adopted as suitable learning or searching procedures for the solution of artificial problems. The first example of an algorithm modelling natural systems was the so-called Genetic Algorithm. The algorithm was based on the concepts of natural selection and genetic pressure. Its implementation was inspired by the studies of Darwin and Mendel on the higher possibility for an individual or "agent" that better fit the surrounding environment to generate offspring and preserve its genetic features throughout successive generations. The success obtained by this optimization approach was immediate and it received a wide and rapid diffusion. Different versions of the original binary GA were developed for the optimization of functionals [47] and they have been applied to real-world applications. The first text book on GAs and related applications was published by Goldberg [69] in 1989.

As compared to previous optimization algorithms, the GA showed many interesting features. More specifically, (a) the GA does not require neither the analytical knowledge

nor the differentiation of the functions to be optimized, but only the values of the fitness are enough to pursue the evolutionary process, (b) the algorithm tends to move towards the most attractive region of the solution space by means of an "almost" blind search technique since the operators are applied in a probabilistic way instead of considering definite rules, (c) sampling the search space not in a single point but in several locations at each iteration and the way the operators recombine the information coded in the population of solutions foster the global search capability of the optimization.

Besides the explicit parallelism guaranteed by its multiple-agent nature as for other population-based stochastic algorithms successively developed, the *GA*s is also related to the concept of schemata (i.e., the building blocks coding each trial solution) and the "implicit parallelism". In [75] and [69], it has been shown that the effective number of schemata [69] processed by the *GA* at each iteration of the evolutionary process is greater than the number of individuals P of the population. Such a property guarantees that, also in a serial implementation, multiple characteristics (i.e., the corresponding schemata) of the solution are processed in parallel. A well-known result is the Holland's inequality stating a lower bound on the order of $P^3/\epsilon\sqrt{l}$ to the number of schemata processed in a population of $P = \xi 2^l$ strings, ξ and l being a small integer [17] and the string length of binary digits, respectively. This result has been generalized in [11] for a population of $P = 2^{\beta l}$ individuals by proving that the schemata bound is a monotonically decreasing function of β and that when $\beta > 4/3$, the expected number of processed schemata is a constant and its lower bound is proportional to $P^{(2log_2 3)/\beta}/\sqrt{log_2 P}$.

After the diffusion of *GA*-based algorithms, the development of artificial systems based on the concepts of swarm intelligence has been more recently considered [16] and new implementations of innovative metaheuristics exploiting the cooperation paradigm, instead of the competitive one of the *GA*s, have been proposed. In this framework, the Particle Swarm Optimizer [87] and the Ant Colony Optimizer [54] have been successfully applied to an increasing number of problems and applications. These algorithms artificially model the social interaction and cooperation of swarm of bees or colony of ants. Accordingly, the activity of each agent is guided not only by the work in progress (namely, its fitness to the environment) but also taking into account the information coming from the interactions with other agents or present in the local environment (i.e., the stigmergy ‡).

‡ The concept of stigmergy, whose meaning and implications will be better specified in the following, is widespread and quite important in artificial intelligence. As a matter of fact, it is related to the self-

3. Evolutionary Algorithms - General Framework

*EA*s are iterative procedures, where a pool of P solutions, $F = \left\{\underline{f}^{(p)};\, p = 1, ..., P\right\}$, evolves to find the solution of the problem at hand through the optimization of a suitable function $\Phi\left(\underline{f}^{(p)}\right)$ or functions $\Phi_t\left(\underline{f}^{(p)}\right)$, $t = 1, ..., T$, (multi-objective optimization [43]) aimed at measuring the "goodness" of the trial solution under given constraints. The cost function is the unique link between the optimization problem and the physical one and great attention should be paid to the define Φ in order to obtain representative and reliable solutions at the end of the optimization process. Moreover, the complexity of the cost function as well as the computational burden of its minimization strongly influence the use of a class of optimization algorithms rather than others.

As far as the the design of an *EA*-based optimization technique is concerned, the key points to be addressed are:

- the representation of the solution, $\underline{f} = \{f_n;\, n = 1, ..., N\}$, coding a set of N parameters f_n, $n = 1, ..., N$, to be optimized;
- the design of the evolutionary operators, $\mathcal{L}$, for generating the succession (ideally infinite) of trial solutions, $\underline{f}_k$, $k = 0, ..., \infty$, k being the iteration index;
- the evolution procedure, namely the criteria and guidelines to generate new solutions by means of the evolutionary operators.

At the initialization of the iterative process, the initial set of solutions $F_0 = \left\{\underline{f}_0^{(p)};\, p = 1, ..., P\right\}$ is usually randomly-generated within the search space

$$f_n = r f_n^{max} + (1 - r)\, f_n^{min} \tag{1}$$

starting from the knowledge of the upper f_n^{max} and lower f_n^{min} bounds of the parameter f_n that limit the admissible search space Ω (i.e., $\underline{f} \in \Omega$ with $f_n \in [f_n^{min},\, f_n^{max}]$). Moreover, $r \in [0, 1]$ is a uniformly-distributed random variable. Otherwise, the initial population can be defined on the basis of some *a-priori* information on the problem at hand and its solutions. In such a case, the solutions are statistically-generated around a reference trial solution $\underline{\hat{f}} = \left\{\hat{f}_n;\, n = 1, ..., N\right\}$ by considering either a uniform

$$f_n = \hat{f}_n \left(\frac{2r - 1}{2}\right) \tag{2}$$

organization process [16]. In the framework of optimization, it means that the surrounding architecture (i.e., the environment/solution space) provides a sufficient amount of information and constraints to control the low level actions (i.e., those of the single agents) such that the general activity of the entire swarm/colony seems being governed by a global plan. The notion has been first introduced by Grassé in 1959 about the termites' behavior [71] and an interesting review on the subject can be found in [155].

or a normal distribution

$$G(f_n) = \frac{1}{\sqrt{2\pi}\varsigma_n} \exp - \left(\frac{f_n - \hat{f}_n}{\sqrt{2}\varsigma_n} \right)^2 \tag{3}$$

ς_n being a real index controlling the statistical distribution of the parameter values.

Successively ($k \geq 1$), a sequence of trial solutions is generated by applying the operators $\mathcal{L}$ in a stochastic fashion and according to the adopted evolutionary procedure. The pool of solutions at the $(k+1)$-th iteration, F_{k+1}, is given by

$$\underline{f}_{k+1}^{(p)} = \underline{f}_{k}^{(p)} + \underline{s}_{k+1}^{(p)}\,,\ p = 1, ..., P \tag{4}$$

where $\underline{s}_{k+1}^{(p)}$ is defined on the basis of the solutions F_k at the previous iteration through the application of the evolutionary operators

$$\underline{s}_{k+1}^{(p)} = \mathcal{L}\{F_k\}\ . \tag{5}$$

Accordingly, a trial solution at iteration $k+1$ turns out to be expressed

$$\underline{f}_{k+1}^{(p)} = \underline{f}_{0}^{(p)} + \sum_{j=0}^{k} \mathcal{L}\{F_j\} \tag{6}$$

where

$$\begin{aligned} \mathcal{L}\{F_k\} &= \mathcal{L}\left\{ \underline{f}_{k}^{(p)};\ p = 1, ..., P \right\} \\ &= \mathcal{L}\left\{ \underline{f}_{k-1}^{(p)} + \underline{s}_{k}^{(p)};\ p = 1, ..., P \right\} \\ &= \mathcal{L}\left\{ \underline{f}_{k-1}^{(p)} + \mathcal{L}\{F_{k-1}\};\ p = 1, ..., P \right\}\ . \end{aligned} \tag{7}$$

The structure of the *EA*s is then fully described by detailing the following architecture levels, namely the "Basic level" and the "Control level".

3.1. Basic Level

The *basic level* is responsible for the generation of the succession of trial solutions and it is concerned with the coding of the solutions and the design of the evolutionary operators.

The coding of the problem unknowns, f_n, $n = 1, ..., N$, through a set of symbols belonging to an alphabet $\mathcal{A}$ is a key point of the *EA*s since it forces the choice of the evolutionary operators as well as the granularity of the optimization and the accuracy of the final solution. The most popular coding strategies, widely used in several practice

applications, are the binary coding, $\mathcal{A} = \{0, 1\}$, and the real coding, $\mathcal{A} \equiv \{\mathbb{R}\}$. The easy implementation of the former in personal computers and the fact that many problems deal with continuous real-valued variables have contributed to the proliferation of EAs with these coding strategies as well as the design of customized real/binary evolutionary operators. On the other hand, the use of a discrete alphabet of S symbols, $\mathcal{A} = \{a_1, ..., a_S\}$, is common in combinatorial optimization.

Generally speaking and whatever the alphabet, a coding law $\Gamma(\cdot)$ is used to map the set of parameters, $\underline{f} = \{f_1, ..., f_N\}$, from the input space (called *phenotype space*) to its coded representation, $\underline{c} = \{c_1, ..., c_M\}$, in the work space (called *genotype space*): $\underline{c} = \Gamma\left(\underline{f}\right)$, $M \geq N$. Although the terms phenotype and genotype come from genetics and were first introduced by Holland [75] in dealing with artificial adaptive systems, their meaning is more general and it is not limited to the framework of genetic-based optimization algorithms. As regards to the coding function, it can be defined between equal-dimensional spaces (i.e., $M = N$) or to a higher dimensionality (i.e., $M > N$). Once a new set of coded solutions is determined in the genotype space by using the evolutionary operators, a decoding law is applied to map the updated coded parameters into a new trial solution within the phenotype space: $\underline{f} = \Gamma^{-1}(\underline{c})$.

Concerning the evolutionary operators, they are usually inspired by natural paradigms. Representative examples are those modeled on the concepts of natural selection (GAs and DE), cooperation and stigmergy taken from the intelligence of swarms (e.g., PSO and ACO), and distribution of knowledge [e.g., Memetic Algorithm (MA)].

3.2. Control Level

The *control level* is the architectural structure devoted to exploit the building blocks of the basic level in sampling the solution space to find the global optimum. At this level, the issues related to the setup of the control parameters, the definition of the termination conditions, and the introduction of the problem constraints [e.g., $h_i\left(\underline{f}_k^{(p)}\right) = 0$, $i = 1, ..., I$, or $g_j\left(\underline{f}_k^{(p)}\right) \leq 0$, $j = 1, ..., J$] or boundary conditions (e.g., $f_n \in [f_n^{min}, f_n^{max}]$ such that $\underline{f} \in \Omega$) on the solutions are properly addressed. More specifically, the control parameters define the number of agents or dimension of the population/swarm of trial solutions, P_k, used at each iteration and the probabilities of applying the evolutionary operators $\mathcal{L}$. As regards to the convergence criteria, simpler termination conditions are based on heuristic assumptions and user-defined thresholds

on the value of the function to optimize or on a maximum amount of iterations, K [69][151][87][54]. More sophisticated choices take into account the stationariness of the optimal cost function value, $\Phi_k^{opt} = \left\{ min_{p=1,...,P} \left[\Phi \left(\underline{f}_k^{(p)} \right) \right] \right\}$, in a fixed range of iterations, K_{window},

$$\frac{\left| K_{window} \Phi_{k-1}^{opt} - \sum_{i=1}^{K_{window}} \Phi_i^{opt} \right|}{\Phi_k^{opt}} \leq \eta \tag{8}$$

η being a numerical threshold. Furthermore, conditions quantifying the "diversity" of the solutions of the population are also used [5].

The boundary conditions are usually related to the physical admissibility of the solution and derive from the *a-priori* information on the actual solution. Such an information allows one to reduce the dimension of the search space and is of fundamental importance for the (fast) convergence towards the global optimum.

3.3. Single vs. Multiple Objective Optimization

In single-objective problems (SOPs), the optimization is aimed at looking for the minimum (or maximum) of a scalar function $\Phi \left(\underline{f} \right) : \Omega \subseteq \mathbb{R}^N \rightarrow \mathbb{R}$ subject to some constraints. The solution minimizing the objective function is called global minimum $\underline{f}^{opt} = arg \left\{ min_{\underline{f}} \left[\Phi \left(\underline{f} \right) \right] \right\}$. The sufficient condition for a point of the solution space, $\underline{f} \in \Omega$, to be the global minimum on Ω is that

$$\Phi \left(\underline{f}^{opt} \right) \leq \Phi \left(\underline{f} \right), \forall \underline{f} \in \Omega \,. \tag{9}$$

Dual considerations hold true for maximization problems.

Differently, several problems are mathematically described in terms of vectorial cost function, $\underline{\Phi} \left(\underline{f} \right) : \Omega \subseteq \mathbb{R}^N \rightarrow \mathbb{R}^T$,

$$\underline{\Phi} \left(\underline{f} \right) = \left[\Phi_1 \left(\underline{f} \right), \Phi_2 \left(\underline{f} \right), ..., \Phi_T \left(\underline{f} \right) \right] \tag{10}$$

where each scalar function $\Phi_t \left(\underline{f} \right)$ models a different objective or performance criterion usually conflicting with the others. This is the case of multi-objective problems (MOPs) or vector optimization problems [43][142][143][175] dealing with multicriteria optimization. Unlike SOPs, the meaning of optimum modifies into the "best" trade-off solution among the whole set of performance criteria. The notion of Pareto optimality [113] is generally adopted to properly model this concept. A solution $\underline{f}$ is Pareto-optimal

on Ω if no other solutions exist that dominate it. A solution $\underline{f}^{(a)}$ (strictly) dominates $\underline{f}^{(b)}$ if and only if

$$\Phi_t\left(\underline{f}^{(a)}\right) \leq \Phi_t\left(\underline{f}^{(b)}\right) \ , \ t = 1, ..., T \tag{11}$$

and

$$\exists t \in [1, T] : \ \Phi_t\left(\underline{f}^{(a)}\right) < \Phi_t\left(\underline{f}^{(b)}\right) . \tag{12}$$

As a consequence, (a) Pareto optimal solutions cannot reduce their performances on a criterion $\Phi_{t'}$ without increasing their effectiveness in fitting at least another criterion $\Phi_{t''}$; (b) the solution of a MOP is not unique, but all Pareto optimal solutions are suitable solutions; (c) the solutions on Ω non-strictly dominated generate the so-called Pareto front. Since, no general rules exist for the choice of the best solution in MOPs, the global optimum is chosen either according to the user-requirements or by reformulating the MOP into an equivalent SOP whose scalar cost function is the linear combination of the MOP objective functions

$$\Phi\left(\underline{f}\right) = \sum_{t=1}^{T} w_t \Phi_t\left(\underline{f}\right) \tag{13}$$

w_t, $t = 1, ..., T$, being real user-defined coefficients. As far as solution algorithms for the MOPs are concerned, although many mathematical programming procedures have been designed for the retrieval of the solutions of the Pareto front [92], EAs seems to be very suitable to MOPs because of their intrinsic/implicit parallelism that allows to simultaneously manage a set of different solutions [43] and to find multiple solutions at each iteration. Moreover, EAs can easily address optimization problems whose Pareto fronts are either discontinuous or concave while the searching capabilities of other optimizers turn out to be more dependent on the nature of the Pareto front.

4. Evolutionary Algorithms - Implementations

In this section, a brief overview on EAs usually (to the best of the authors' knowledge) applied to the solution of inverse scattering problems is reported. The section is subdivided in three main parts. The first one is devoted to describe genetic-based optimization algorithms. Standard implementation of GAs and DE are presented pointing out the main differences and common features. Unlike GAs and DE whose underlying architecture models a competitive and hierarchical framework aimed at

promoting the reproduction/evolution of fittest individuals, decentralized optimization procedures based on the intelligence of swarms, namely the particle swarm optimizer and the ant colony optimizer, are considered in the second part. Finally, some state-of-the-art hybrid algorithms are briefly summarized.

4.1. Genetic-based Optimization

4.1.1. Genetic Algorithms - *GA*s are *EA*s modeled on concepts of natural selection and genetic pressure to perform an effective sampling of the solution space. *GA*s basic principles were first introduced by Holland in 1975 [75] and extended to functional optimization by De Jong [47] with an immediate diffusion to real-world problems because of their effectiveness in dealing with complex functions [69][77][61] when compared to standard deterministic procedures.

The solutions at the k-th iteration and belonging to the phenotype space, $\underline{f}_k^{(p)}$, $p = 1, ..., P$, are called *individuals*, while their corresponding version in the genotype space are denoted as *chromosomes*, $\underline{c}_k^{(p)}$, $p = 1, ..., P$. At each iteration or *generation*, the set $F_k = \left\{ \underline{f}_k^{(p)}; \, p = 1, ..., P \right\}$ of P agents or individuals compose a *population* of trial solutions named *parents*. The set of genetic operators $\mathcal{L}^{GA}$ is applied to F_k to generate a new population F_{k+1}. More specifically, the selection, $\mathcal{S}$, the crossover, C, and the mutation, $\mathcal{M}$, act on the parents to determine the individuals of the new population, called *children* or *offspring*.

In their basic version, the *GA*s follow the workflow in Fig. 1. After the initialization of the population F_0 at the first iteration ($k = 0$), new populations, F_k, $k \geq 1$, are iteratively generated by applying the genetic operators $\mathcal{L}^{GA} = \{\mathcal{S}, C, \mathcal{M}\}$ as follows. For each iteration, a "mating pool" is chosen by applying the selection procedure to F_k

$$F_{k(\mathcal{S})} = \mathcal{S}\left\{F_k\right\} \, . \tag{14}$$

Standard implementations consider the roulette-wheel selection or the tournament selection [69]. The selection procedure performs taking into account the knowledge on how current individuals fit the problem at hand. Mathematically, such a knowledge is acquired by computing the cost function values of the current population, $\Phi_k^{(p)} = \Phi\left(\underline{f}_k^{(p)}\right)$, $p = 1, ..., P$. Fittest individuals have higher probability to be chosen as parents for generating new individuals and for reproducing their chromosomes [75]. Dealing with minimization problems and according to a fitness proportional selection mechanism, the

probability of a parent to be chosen for the "mating pool" is equal to

$$\lambda_{k(\mathcal{S})}^{(p)} = \frac{1/\Phi_k^{(p)}}{\sum_{i=1}^{P} 1/\Phi_k^{(i)}} \,. \tag{15}$$

A new population is then generated by applying crossover and mutation according to the values of the probabilistic coefficients defined at the control level

$$F_{k+1} = F_{k(\mathcal{C})} \cup F_{k(\mathcal{M})} \tag{16}$$

where $F_{k(\mathcal{C})}$ and $F_{k(\mathcal{M})}$ indicate the set of new individuals obtained by crossover

$$F_{k(\mathcal{C})} = \mathcal{C}\left\{F_{k(\mathcal{S})}\right\} \tag{17}$$

and mutation

$$F_{k(\mathcal{M})} = \mathcal{M}\left\{F_{k(\mathcal{S})}\right\} \,, \tag{18}$$

respectively. The genetic operators are iteratively applied on the mating pool until the population is completed and the parents, whom neither crossover nor mutation are applied to, are directly reproduced in the next population. To enhance the convergence behavior of *GA*s, another operator known as *elitism* is often used. The elitist strategy is applied whether the condition $\Phi_{k+1}^{opt} > \Phi_k^{opt}$ (in minimization problems) holds true and it consists in inserting the best individual of the k-th iteration in place of the worst solution of the successive iteration.

To further improve the convergence as well as the global capability of *GA*s, besides the commonly-used genetic operators, the *GA*s have been also modified by using enhanced techniques like dominance and diploidy, sharing, or knowledge-based operators [69].

(A) Binary *GA*s

Genetic Algorithms were firstly implemented to work with binary or discrete unknowns. The problem unknowns are coded, if not already binary, in strings of $l = \sum_{n=1}^{N} Q_n$ bits, Q_n is the number of levels used to quantize the range of existence of the n-th unknown parameter. Both uniform [90]

$$\tilde{f}_n = f_n^{min} + \left[\frac{f_n^{max} - f_n^{min}}{2^{Q_n} - 1}\right] \sum_{i=0}^{Q_n - 1} a_{i(n)} 2^i \tag{19}$$

or non-uniform quantization [76]

$$\tilde{f}_n = \sum_{i=1}^{Q_n} a_{i(n)} 2^{1-i} \chi_n \tag{20}$$

can be used. More in detail, where $a_{i(n)}$, $i = 1, ..., Q_n$, is the set of bits (or *alleles*) composing the coded parameter, $c_n = \left\{a_{1(n)}, ..., a_{Q_n(n)}\right\}$, and $\chi_n = \frac{f_n^{max}}{2}$ is the largest quantization level.

Genetic operators $\mathcal{L}^{GA}$ act on the chromosomes $\underline{c}^{(p)}$, $p = 1, ..., P$ as follows. In the selection phase, a pair of parents $\underline{c}_k^{(p_1)}$ and $\underline{c}_k^{(p_2)}$ is chosen. The recombination of the strings of genes is then performed through crossover with probability p_C. By considering the single-point crossover, an integer value $i \in [1 : l]$ is randomly chosen and two children are generated whose chromosomes turn out being equal to $\underline{c}_{k+1}^{(p_1)} = \left\{a_{1,k}^{(p_1)}, ...a_{i,k}^{(p_1)}, a_{i+1,k}^{(p_2)}, ..., a_{B,k}^{(p_2)}\right\}$ and $\underline{c}_{k+1}^{(p_2)} = \left\{a_{1,k}^{(p_2)}, ...a_{i,k}^{(p_2)}, a_{i+1,k}^{(p_1)}, ..., a_{B,k}^{(p_1)}\right\}$. Each child contains parts of the genetic structure of both parents. Moreover, an individual is mutated with probability p_M by randomly flipping the value from 1 to 0 or viceversa of some alleles of the corresponding chromosome, p_{MB} being the bit mutation probability. Obviously, more complex implementations of the crossover operator (e.g., two-point crossover, uniform crossover [161]) and the mutation (e.g., interchange [90]) exist, as well.

Since the binary GA (BGA) works with a finite dimension parameter space, it turns out to be more adapt to deal with problems where the unknowns can assume only a finite number of values. Concerning real (continuous) variables, unknown parameters need to be quantized with an unavoidable quantization error. This error can be reduced by increasing the gene length l at the cost of a decrease in the convergence speed and an increase in the memory requirements. Moreover, the GA operators acting on a binary-coded representation of the solution do not assure that the chromosomes of the next generation are admissible solutions. Moreover, if acceptable solutions have to belong to some domains of the solution space (e.g., when constraints are imposed), monitoring this property under the action of the genetic operators can be laborious and time-consuming (a decoding should be performed), and the convergence may therefore be slowed.

(B) Real coded GAs

Binary encoding is not the only way to represent a parameter when applying GAs. In the presence of real parameters, it is more logical to use the floating point representation [89]. For the real-coded GA (RGA), a gene is represented by the value of the unknown itself

$$\underline{c}_k^{(p)} = \underline{f}_k^{(p)} \,. \tag{21}$$

As a consequence, new genetic operators are designed although the peculiarities of the original operators should be maintained. Mutation and crossover must remain a mean to explore the parameter space randomizing selected solutions and a way to (randomly) mix the good characteristics of the chromosomes, respectively.

In *RGA*s, the crossover is defined as the arithmetical linear combination of two chromosomes. Once two parents $\underline{f}_k^{(p_1)}$ and $\underline{f}_k^{(p_2)}$ are selected, the resulting offspring are given by

$$\begin{aligned} f_{n,k+1}^{(p_1)} &= r f_{n,k}^{(p_1)} + (1-r) f_{n,k}^{(p_2)} \\ f_{n,k+1}^{(p_2)} &= (1-r) f_{n,k}^{(p_1)} + r f_{n,k}^{(p_2)} \quad , \quad n = 1, ..., N. \end{aligned} \tag{22}$$

The *RGA* mutation consists in adding a random value $s_n \in \left\{ f_{n,k}^{(p)} - f_n^{min},\ f_n^{max} - f_{n,k}^{(p)} \right\}$ to a randomly-selected pth chromosome

$$f_{n,k+1}^{(p)} = f_{n,k}^{(p)} + s_n \,. \tag{23}$$

Whether the new trial solutions are not physical and do not belong to the solution space ($\underline{f} \notin \Omega$), they are modified exploiting the *a-priori* knowledge on the boundaries of the solution space as follows

$$f_{n,k+1}^{(p)} = \begin{cases} f_n^{min} & if\ f_{n,k}^{(p)} < f_n^{min} \\ f_n^{max} & if\ f_{n,k}^{(p)} > f_n^{max} \end{cases} . \tag{24}$$

The *RGA* gained increasing popularity because it is easy to implement, computationally efficient when dealing with a small number of real-valued unknowns, and suitable for fine-tuning the selective pressure [4].

(C) Hybrid coded *GA*s

In some problems, the *a-priori* knowledge on the solution allows a parametrization of a subset of the unknowns through a small set of discrete descriptors

$$f_d = \mathcal{H}\left\{ d_j;\ j = 1, ..., J \right\},\ d \in [1, L] \tag{25}$$

where $\{d_j;\ j = 1, ..., J\}$ is the sub-set of discrete *equivalent* parameters being $J < L < N$. In such a case, a suitable encoding procedure must be defined in order to provide a one-to-one mapping between the phenotype space and the genotype space, but at the same time, exploiting the features of the unknown parameter set. A hybrid coded *GA*

(HGA) is described in [22] to deal with microwave imaging problems. A set of integer-valued equivalent parameters, $\{d_j;\, j=1,...,J\}$, is binary coded and a floating-point representation is used for the remaining real unknowns, $\{f_n;\, n=L+1,...,N\}$. As far as the genetic operators are concerned, they have been properly modified to maintain the structure of the hybrid chromosomes. During mutation, if the gene to be perturbed is binary-coded, it is changed from 0 to 1 or viceversa as for the BGA. Otherwise, the mutation (23) defined for the RGA is considered.

Concerning the efficiency and effectiveness of HGA-based strategies, the exploitation of some a-priori information to define a suitable parametrization function $\mathcal{H}$ and the choice of a reduced set, instead of the whole number, of representative parameters is of fundamental importance to reduce the dimension of the search space. Moreover, the parametrization method can be profitably adopted to prevent the generation of solutions physically not admissible.

4.1.2. Differential Evolution - Unlike GAs, the Differential evolution algorithm has been originally proposed by Storn and Price [150] for the global optimization over continuous spaces. They were mainly aimed at simplifying the evolution process of GAs as well as to enhance the convergence rate [151][121]. The iterative evolution of the DE is similar to that of the GAs. Each current population is replaced by better individuals obtained by applying the DE operators, $\mathcal{L}^{DE}$, still based on genetic but now executed in a different sequence: the mutation, $\mathcal{M}$, the crossover, C, and selection, $\mathcal{S}$.

The DE iteratively evolves as shown in Fig. 2. During the mutation process, an intermediate solution is generated in correspondence with each individual $\underline{f}_k^{(p)}$ as follows

$$\underline{t}_{k+1}^{(p)} = \underline{f}_k^{(p_1)} + \varepsilon \left(\underline{f}_k^{(p_2)} - \underline{f}_k^{(p_3)} \right) \;,\; p=1,...,P \tag{26}$$

where p, p_1, p_2, $p_3 \in [1,P]$ ($p \neq p_1 \neq p_2 \neq p_3$) are the indexes of different individuals randomly chosen in F_k. The agents $\underline{f}_k^{(p_1)}$, $\underline{f}_k^{(p_2)}$, and $\underline{f}_k^{(p_3)}$ are called *donor* vectors or *secondary* parents, and $0 < \varepsilon \leq 2$ is a real and constant value that controls the amplification of the differential variation $\left(\underline{f}_k^{(p_2)} - \underline{f}_k^{(p_3)} \right)$. The crossover is then applied between the intermediate solution, $\underline{t}_{k+1}^{(p)}$, called *mutant* vector and the *primary* parent, $\underline{f}_k^{(p)}$, according to the following strategy

$$\underline{u}_{k+1}^{(p)} = \begin{cases} \underline{t}_{k+1}^{(p)} & if \;\; (r < p_C) \\ \underline{f}_k^{(p)} & otherwise \end{cases} \;. \tag{27}$$

Finally, the selection takes place and $\underline{f}_{k+1}^{(p)}$ is chosen according to a greedy criterion by comparing $\Phi\left(\underline{f}_{k}^{(p)}\right)$ with $\Phi\left(\underline{u}_{k+1}^{(p)}\right)$. In a minimization problem, $\underline{f}_{k+1}^{(p)} = \underline{u}_{k+1}^{(p)}$ when $\Phi\left(\underline{u}_{k+1}^{(p)}\right) \leq \Phi\left(\underline{f}_{k}^{(p)}\right)$, while $\underline{f}_{k+1}^{(p)} = \underline{f}_{k}^{(p)}$ otherwise.

As regards to the control parameters of DE, they are the crossover probability p_C and the amplification coefficient ε to be carefully chosen to avoid a premature convergence to sub-optimal solutions or a slow convergence rate [151].

As compared to GAs, the main differences are (a) the order of execution of the genetic operators (Fig. 2) and (b) the competition between parents and children during the selection phase which lacks in GAs since the offspring are all accepted while the parents are all discarded. Unlike GAs, the fittest parents have higher probability to generate children with better fitness. Moreover, the risk that the average fitness of the population can get worse is greater in GAs since crossover and mutation are performed after selection. Furthermore, since the secondary parents are chosen from the population with equal probability (and not through a proportional fitness selection), the DE usually increases its global searching capabilities. Finally, the cost function of the best individual Φ_k^{opt}, $k = 1, ..., K$, monotonically decreases in the DE because of the DE implementation of the selection mechanism and without the need of particular elitist strategy.

Despite the basic version of the DE, many different versions of the algorithm exist [1][121][33]. To identify them the notation $DE/x/y/z$ is generally adopted [151]. More in detail,

- x is the mutated solution randomly chosen ($x = rand$) or set to the best individual within the population ($x = best$);
- y is equal to the number of difference vectors used in the differential variation;
- z indicates the crossover scheme.

According to such a notation, the version of the DE presented above is identified as "$DE/rand/1/bin$". For completeness, let us notice that the mutation operator of the version "$DE/best/2/bin$" is defined as

$$\underline{t}_{k+1}^{(p)} = \underline{f}_{k}^{opt} + \varepsilon\left(\underline{f}_{k}^{(p_1)} + \underline{f}_{k}^{(p_2)} - \underline{f}_{k}^{(p_3)} - \underline{f}_{k}^{(p_4)}\right) . \tag{28}$$

4.2. Optimization by Swarm Intelligence

4.2.1. Particle Swarm Optimizer - The particle swarm optimizer is a robust stochastic search procedure suitable for the optimization of contiguous unknowns inspired by the

social behavior of insect swarms, school of fishes and flocks of birds. In the PSO, an agent, $\underline{b}_k^{(p)}$, called *particle* is characterized by a position $\underline{f}_k^{(p)}$ in the solution space and a velocity $\underline{v}_k^{(p)}$ that models the capability of the p-th particle to fly from the current position to another successive position $\underline{f}_{k+1}^{(p)}$. The whole set of particles $\left\{\underline{b}_k^{(p)}, p = 1, ..., P\right\}$, constitutes the *swarm* F_k. In its classical implementation [85], the particle update equations are

$$\underline{f}_{k+1}^{(p)} = \underline{f}_k^{(p)} + \underline{v}_{k+1}^{(p)} \tag{29}$$

and

$$v_{n,k+1}^{(p)} = \omega v_{n,k}^{(p)} + C_1 r_1 \left(p_{n,k}^{(p)} - f_{n,k}^{(p)}\right) + C_2 r_2 \left(g_{n,k} - f_{n,k}^{(p)}\right) \tag{30}$$

whose physical interpretation, derived by Newton's laws, has been given in [107]. In (30), ω, C_1, and C_2 are control parameters known as *inertial weight*, *cognitive* and *social* acceleration terms, respectively [87]. Moreover, r_1 and r_2 are two random variables having uniform distribution in $[0, 1]$. With reference to a minimization problem, the values $\underline{p}_k^{(p)} = arg\left\{min_{i=1,...,k}\left[\Phi\left(\underline{f}_i^{(p)}\right)\right]\right\}$ and $\underline{g}_k = arg\left\{\underset{i=1,...,k;\, p=1,...,P}{min}\left[\Phi\left(\underline{f}_i^{(p)}\right)\right]\right\}$ are the so-called *personal* and *global* best solutions, namely the best positions found by the p-th particle and by the whole swarm until iteration k, respectively.

As far as the iterative optimization is concerned (Fig. 3), starting from guess values of $\underline{f}_0^{(p)}$ and $\underline{v}_0^{(p)}$, $p = 1, ..., P$, the positions and velocities of the particles are updated according to Eqs. (29) and (30).

The main advantages of the PSO if compared to either the GAs or the DE can be summarized in the followings:

- the simplicity of the algorithm implementation and the use of a single operator (i.e., the velocity update) instead of three genetic operators (i.e., the crossover, the mutation, and the selection);
- the easy manipulation of the calibration parameters [138] (i.e., the swarm size, the inertial weight, and the acceleration coefficients) which control the velocity update operator. Even if the number of control parameters (i.e., the population size, the crossover rate, the mutation rate) is similar, it is certainly easier to set the PSO indexes than evaluating the optimal setting among various operators and several options of implementation;

- the ability to prevent the stagnation by controlling the inertial weight and the acceleration coefficients to sample new regions of the solution space. In standard *GA*s and *DE*, the stagnation occurs when the trial solutions assume the same genetic code close to that of the fittest individual. In such a case, the crossover does not contribute to the evolution and only a lucky mutation could locate a new individual in other interesting region of the solution space;
- a smaller number of agents, which turns out in a reduced computational cost of the overall optimization and enables a reasonable compromise between computational burden and efficiency of the iterative process.

Although the *PSO* is intrinsically an optimizer for continuous spaces, a binary version of the algorithm exists [86], as well. In order to deal with discrete spaces [86], the concepts of trajectory, position and velocity have been properly redefined in terms of changes of probabilities. More specifically, each dimension of the solution space is normalized to assume values between 0 and 1. In such a space, the velocity is constrained to the same range of variation and its value gives the probability threshold for having a binary allele with zero or one value. The new allele is then computed as follows

$$f_{n,k+1}^{(p)} = \begin{cases} 1 & if\ r < \Lambda\left(v_{n,k+1}^{(p)}\right) \\ 0 & otherwise \end{cases} \tag{31}$$

by defining a suitable transformation function $\Lambda\left(v_{n,k+1}^{(p)}\right)$ usually consisting of a sigmoid

$$\Lambda\left(v_{n,k+1}^{(p)}\right) = \frac{1}{1 + exp\left(-v_{n,k+1}^{(p)}\right)}. \tag{32}$$

4.2.2. Ant Colony Optimizer - The *ACO* is a population-based global optimization algorithm inspired by the foraging behavior of ant colonies looking for food sources [52]. The ants move in the space surrounding the nest looking for the best (shortest) path between the food sources and the nest. Likewise the *PSO*, the *ACO* is based on the concepts of swarm intelligence and cooperation, but it also exploits the paradigm of stigmergy and self-organization [64]. In this sense, the activity of each agent $\underline{f}_k^{(p)}$ or *ant* in the *colony* F_k is guided not only by the work in progress (the goal of optimization), but also from the information available in the local environment. To modify the local environment, each ant leaves a chemical substance called *pheromone* while moving within the solution space along a path. The amount of pheromone on a path quantifies its degree of optimality, but it decays with time (*evaporation* mechanism). These

mechanisms allow one to avoid poor food sources on one hand and on the other to efficiently sample the whole solution space.

The first implementation of the *ACO* [52] was originally developed for discrete optimization problems and it was applied to solve complex combinatorial problems [53][95]. In its basic version (Fig. 4) concerned with the search of a path within a discrete space (e.g., in the Traveling Salesman Problem [53]), each ant codes a vector $\underline{f}_k^{(p)}$ representative of a set of discrete symbols or locations, $\underline{f}_k^{(p)} = \{a_1, ..., a_N\}$. Let us suppose ψ_k^{ij} be the amount of pheromone on the edge between the location a_i and a_j, $i, j \in [1, N]$, with $i \neq j$. Every vector, $\underline{f}_k^{(p)}$, $p = 1, ..., P$, is randomly initialized at the first iteration ($k = 0$) and a uniform level of pheromone is assigned to each path within the search space, $\psi_0^{ij} = cost$. Successively, the pheromone level of each edge of the path covered by the p-th ant is updated

$$\tilde{\psi}_{k+1}^{ij} = \psi_k^{ij} + \sum_{i=1}^{P} \delta\left\{\psi_k^{ij}, \underline{f}_k^{(p)}\right\} \frac{H}{\Phi\left(\underline{f}_k^{(p)}\right)}, \ \forall \psi_k^{ij} \tag{33}$$

where $\delta\left\{\psi_k^{ij}, \underline{f}_k^{(p)}\right\} = 1$ when $\psi_k^{ij} \in \underline{f}_k^{(p)}$ (i.e., the path crossed by solution $\underline{f}_k^{(p)}$ contains the branch individuated by ψ_k^{ij}) and $\delta\left\{\psi_k^{ij}, \underline{f}_k^{(p)}\right\} = 0$, otherwise. Moreover, H is a real positive constant term. The evaporation procedure takes place to reduce the amount of pheromone on each path of the graph

$$\psi_{k+1}^{ij} = (1 - \rho)\, \tilde{\psi}_{k+1}^{ij}, \ \forall \psi_k^{ij} \tag{34}$$

$\rho \in (0\,,\, 1]$ being a parameter aimed at controlling the evaporation rate. At each iteration, the probability to move towards a new position a_j within the graph leaving the position a_i is given by

$$p_\psi\left(\psi_{k+1}^{ij}\right) = \frac{\psi_{k+1}^{ij}}{\sum_j \Xi\left(\psi_{k+1}^{ij}\right)} \tag{35}$$

where $\Xi\left(\psi_{k+1}^{ij}\right) = \psi_{k+1}^{ij}$ if ψ_{k+1}^{ij} is a physically-admissible path and $\Xi\left(\tau_{k+1}^{ij}\right) = 0$, otherwise.

The *ACO* has been also extended to the optimization in continuous spaces [12][148][149]. The Continuous *ACO* (*CACO*) considers a solution archive where $Y > P$ solutions are stored and used to generate new solutions. The value Y depends upon the problem complexity and each solution of the archive is identified by its fitness to the problem at hand, $\Phi_k^{(p)} = \Phi\left(\underline{f}_k^{(p)}\right)$, $p = 1, ..., Y$. At each iteration, a new set of P solutions is probabilistically-generated and added to the solution archive. The $Y + P$

solutions are ranked from the best ($p = 1$) to the worst ($p = Y$) according to the corresponding fitness. Successively, the worst P solutions are then removed form the archive.

The new solutions are obtained sampling a suitable probability density functions, Θ_k. Usually, the probability density function is a weighted sum of N dimensional Gaussian functions

$$\Theta_k\left(f_n\right) = \sum_{p=1}^{Y} w_k^{(p)} \frac{1}{\varsigma_{n,k}^{(p)}\sqrt{2\pi}} exp\left[-\frac{\left(f_n - \overline{f}_n^{(p)}\right)^2}{2\left(\varsigma_{n,k}^{(p)}\right)^2}\right] \;,\; n = 1, ..., N \tag{36}$$

where the mean value $\overline{f}_{n,k}^{(p)}$ and the standard deviation $\varsigma_{n,k}^{(p)}$ are given by $\overline{f}_n^{(p)} = f_{n,k}^{(p)}$, $n = 1, ..., N$, and $\varsigma_{n,k}^{(p)} = \rho \sum_{i=1}^{Y} \frac{\left|a_{n,k}^{(i)} - a_{n,k}^{(p)}\right|}{Y-1}$. Moreover, the weights of the Gaussian functions are defined as

$$w_k^{(p)} = \frac{1}{\varpi Y\sqrt{2\pi}} exp\left[-\frac{(p-1)^2}{2\left(\varpi Y\right)^2}\right] \;,\; p = 1, ..., P \tag{37}$$

ϵ being a parameter modeling a kind of convergence pressure mechanism. When ϖ is small, the best-ranked solutions are preferred, while when it is large, the probability becomes more uniform.

4.3. Hybrid Optimization

Evolutionary algorithms are known as robust optimization techniques able to effectively explore wide parameter spaces. However, *EA*s generally require a high number of cost function evaluations to converge, thus offering reduced performances in terms of computational efficiency when compared to deterministic optimization techniques. However, whether the evaluation of the cost is computationally fast, *EA*s are still very good candidates for a successful solution of the problem at hand especially when local minima are present. Otherwise, when the evaluation of the cost function is cumbersome, different approaches have been proposed to make *EA*-based procedures more competitive still maintaining their positive features. On one hand, suitable encodings (as shown in Sect. 4.1.1) allow a reduction of the dimension of the solution space. On the other hand, to save computational resources and to increase the convergence rate, an effective strategy is the hybridization [14]. As a matter of fact, gradient-based minimization techniques [119] usually converge very fast and yield good results dealing with convex functionals. However, they can be trapped in local minima

in highly nonlinear problems. In order to exploit complementary advantages, EA-based procedures and deterministic methods can be coupled according to the following strategies:

- the EA/CG-based approach;
- the Memetic Algorithm (MA).

4.3.1. EA/CG-based approach - The simplest and more general way to realize a hybridized version of an EA is that of considering a "two-stage optimization". Firstly, the optimization is performed with an EA and subsequently the algorithm switches to a deterministic procedure or viceversa.

In [166], a micro-GA (μGA) has been coupled with a deterministic method proposing a communication criterion for stopping the stochastic algorithm and invoking the deterministic optimizer. Moreover, a hybrid optimization method combining the GA and the Levenberg-Marquardt algorithm (LMA) has been proposed in [169]. The LMA is used to localize a minimum and the minimization procedure switches to the GA in order to climb local minima. Another hybrid procedure based on a RGA-based strategy has been presented in [24] where the global search approach is considered to locate the attraction basin of the global optimum while the CG approach is used to reach the global optimum within the same attraction basin. Whether the convergence threshold is not reached during the deterministic minimization, the RGA restarts with a population whose individuals are randomly-generated around the current optimal solution.

The main drawback of these approaches is the need to evaluate the "quality" of a minimum and/or the closeness of the trial solution to the attraction basin of the global minimum. This requires either an accurate knowledge of the cost function, generally not available, or a heuristic definition of the degree of accuracy of each trial solution.

To overcome the drawbacks, a more sophisticated approach has been presented in [97] which considers a closer coupling between stochastic and deterministic optimizers. The coupling is obtained by means of a step-by-step optimization ($SbS-GA$) where only the best individual of each population undergoes a deterministic optimization for a fixed a limited number of intermediate iterations or until a stationary condition holds true. Successively, standard genetic operators are executed on the whole population of solutions.

4.3.2. Memetic Algorithm - Likewise the step-by-step hybridization, Memetic Algorithms have been introduced to define a closer coupling between stochastic and deterministic optimizers for enhancing the computational efficiency of EAs. Unlike the $SbS-GA$, a stronger coupling between the stochastic approach and the deterministic technique is obtained by introducing a genetic operator which performs a gradient-like based minimization (e.g., "hill-climbing" operator [158], G-Bit improvement [69]).

The MA is a hierarchical algorithm based on the concept of "meme" [46]. A *meme* is a unit of information transmitted when people exchange ideas. Each *idea* is a trial solution $\underline{f}_k^{(p)}$, composed by a set of memes. Since ideas are processed before propagating them, each individual can be assumed as a local minimum/maximum of the cost function Φ. From an algorithmic point of view [109][103], the processing of an idea is simulated by means of a deterministic procedure and its propagation and/or evolution with a stochastic GA-based, or more generally EA-based, technique according to the flowchart shown in Fig. 5. In mode detail, after the initialization ($k=0$) each individual, $\underline{f}_k^{(p)}$, $p=1,...,P$, is considered as initial point for a local optimization procedure in order to obtain a population of local optima, $\hat{F}_{k+1}=\left\{\underline{\hat{f}}_{k+1}^{(p)};\, p=1,...,P\right\}$. Afterwards, an iterative loop is performed where global and local search algorithms are iteratively applied to the whole population until a convergence criterion is satisfied. Furthermore, in order to assure a fast convergence and to preserve the characteristics of the best individual, the elitist strategy is generally adopted.

As compared to other EAs, the MAs exhibit some interesting features. Since the population is only composed of local optima, the individuals move from one minimum to another. Therefore, a limited number of iterations is usually required to converge also with a small population. On the contrary, although very effective in terms of convergence rate, the main drawback of MAs is the unacceptable computational load when the number of unknowns is large and the fact that they need local minimization, where the knowledge of the gradient of the functional is often a must.

5. Evolutionary Algorithms - Theoretical Background

5.1. Convergence Analysis and Control-Parameter Setting

In this section, a theoretical analysis pointing out some interesting issues related to the convergence behavior of EAs is discussed and properly referenced. Some hints on the influence of the various operators on the algorithm behavior as well as some indications

on the values of the control parameters are given.

The study mainly focuses on *GA*s and *PSO* as benchmark algorithms based on different evolutionary concepts: the survival of the fittest and the exploitation of swarm intelligence, respectively.

5.1.1. Genetic Algorithms Several theoretical analyses on the *GA*s behavior are based on the concept of *schemata*, originally introduced by Holland [75] to identify any partial string pattern among those available in the search space that can be processed by the *GA*. In [75], a classical binary *GA*, with reproduction, roulette-wheel selection, single-point crossover and mutation was considered to point out the law for either the growth or decay during the optimization process of some string patterns. In order to illustrate the schemata theorem, let us consider the following example.

With reference to a population with $l = 5$ bits chromosomes, the schemata $*0*11$ is characterized by fixed alleles (i.e., the second, the fourth, and the fifth) and some "don't care" positions (i.e., the first and the third). All possible schemata within a population are expressed in terms of a three letters alphabet $\mathcal{A}^{+} = \{0, 1, *\}$. The total number of admissible schemata is equal to $(2+1)^{l}$, while the number of schemata within a population of P individuals can range from 2^{l} up to $P \times 2^{l}$ since each allele of a chromosome can assume the actual value 0/1 or the don't care symbol. Each schemata is identified by two quantities: the order, $o(\cdot)$, and the length, $\delta(\cdot)$. The schemata order is equal to the number of fixed alleles within the schemata. The length of a schemata is the distance between the first and last position with fixed alleles. For example, the two schemata $\underline{s}^{(t)} = \{0\,1\,*\,*\,1\}$ and $\underline{s}^{(h)} = \{*\,*\,0\,1\,*\}$ have order and length equal to $o\left(\underline{s}^{(t)}\right) = 3$, $o\left(\underline{s}^{(h)}\right) = 2$ and $\delta\left(\underline{s}^{(t)}\right) = 4$, $\delta\left(\underline{s}^{(h)}\right) = 1$, respectively.

The effects of genetic operators $\mathcal{L}^{GA}$ on the survival of a schemata during the evolution of the population have been carefully analyzed in [75] and [69]. Summarizing, the number $m(\underline{s}, k)$ of occurrences of a schemata $\underline{s}$ within a population at the k iteration increases/decreases proportionally to

$$m(\underline{s}, k) \cong m(\underline{s}, k-1) \frac{\Phi_{av}(\underline{s})}{\frac{1}{P}\sum_{p=1}^{P} \Phi_{k-1}^{(p)}} \left[1 - \frac{\delta(\underline{s})}{l-1} p_C - o(\underline{s})\, p_M\right] \tag{38}$$

where $\Phi_{av}(\underline{s})$ is the average fitness of the individuals of the population containing the schemata $\underline{s}$ and the condition $p_M \ll 1$ is assumed.

It is worth pointing out that when crossover and mutation are not used and the individuals directly reproduce throughout the generations only on the basis of the

proportional selection, the effect of replication leads to an exponential growth/decay of schemata having an average fitness above/below the average fitness of the whole population. By supposing

$$\Phi_{av}\left(\underline{s}\right) = \left(1+\iota\right)\frac{1}{P}\sum_{p=1}^{P}\Phi_{k-1}^{(p)} \tag{39}$$

and the percentage ι constant during the optimization process, it turns out that

$$m\left(\underline{s},k\right) \cong m\left(\underline{s},0\right)\left(1+\iota\right)^{k} \tag{40}$$

$m\left(\underline{s},0\right)$ being the occurrence of the schemata $\underline{s}$ within the initial population. Equation (40) points out the exponential effect of the genetic pressure on the convergence of the GA [44]. On the other hand, a schemata survives to crossover and mutation when the following condition holds true

$$\frac{\delta\left(\underline{s}\right)}{l-1}p_C + o\left(\underline{s}\right)p_M < 1 \tag{41}$$

where $\left(1-\frac{\delta(\underline{s})}{l-1}p_C\right)$ and $\left(1-p_M\right)^{o(\underline{s})} \cong 1-o\left(\underline{s}\right)p_M$ since $p_M \ll 1$ are the crossover and mutation survival probabilities, respectively. To satisfy (41), the values usually adopted for the probability of mutation and crossover are $p_C \in [0.5,\, 0.9]$ and $p_M \in [0.001,\, 0.1]$.

The results of such analysis define the so-called *Schemata Theorem* or *Fundamental Theorem* of GAs whose main outcome is that "short, low-order, above-average schemata receive exponentially increasing trials in subsequent generations". In [69], Goldberg also formulated the *Building Blocks Hypothesis* by stating that the GA solution converges to the portion of the search space coded by the building blocks composed by high-fit, short, and low-order schemata which have low probability of being disrupted by crossover and mutation.

Further studies have been successively carried out to give some indications on the convergence of the GAs to the optimal solution Φ^{best}. In [139], a probabilistic analysis on the convergence of a canonical GA is presented. The algorithm is described through a Markov chains model and it is aimed at assessing the converge condition on the sequence of trial solutions

$$lim_{k\to\infty} Pr\left\{\Phi_k^{opt} = \Phi^{best}\right\} = 1\,. \tag{42}$$

By considering a proportional selection mechanism and without elitism, it has been demonstrated that the canonical GA never converges to the global optimum. As a matter of fact, it has been proved that there is a non null probability that, whatever the initial

distribution of the population F_0, the algorithm is able to find a solution with fitness value $\Phi_k^{opt} < \Phi^{best}$, $k \to \infty$. In this sense, it turns out that the Schemata Theorem [75] does not imply the convergence to the global optimum in *static* optimization problems. However, it has been also shown [139] that the elitism can assure global convergence since the "transition time" between whatever two states/solutions of the solution space is finite and the global solution can be found at least one in an unlimited run of the algorithm.

The theoretical analysis carried out by Qi and Palmieri has considered, first separately and then in a unified framework, the effects of the proportional selection, the mutation [122] and the crossover [123] with the assumption of an infinite population size (i.e., $F = \left\{\underline{f}^{(p)}; p = 1, ..., P; P \to \infty\right\}$) over continuous spaces (i.e., $\underline{f} \in \mathbb{R}$). In this sense, the whole solution space is sampled by agents and, thanks to this hypothesis, the distribution of the population can be modeled with a sequence of continuous probability density functions $\Theta(F_k)$, $k = 1, ..., \infty$, instead of using discrete distributions. As far as the genetic operators are concerned, it has been shown in [122] that the selection tends to concentrate the individuals around the fittest solution (i.e., the global optimum) according to the genetic pressure proportional to the value of both the density function $\Theta(F_k)$ and the fitness function $\Phi(F_k)$. Such a mechanism also justifies the effectiveness of *GA*s in dealing with multimodal functionals characterized by multiple global optima. On the opposite, the mutation spreads the distribution obtained after selection proportionally to the convolution between the mutation density and the distribution of the population [122].

Because of the infinite dimension of the population, the selection operator by itself guarantees the convergence to the global solution without the need of mutation. However, the use of mutation is mandatory in real optimization problems when finite populations are used since it enables the exploration of new regions of the solution space.

As regards to crossover (either single-point, multi-point, or uniform [48]), the analysis in [123] shows that it is able to find new solutions in a smarter way as compared to mutation thanks to a good trade-off between exploration and exploitation capabilities. As a matter of fact, the crossover is able to diversify the population. Its iterated application reduces the correlations among the solution parameters while maintaining the marginal distribution of each unknown unaltered and equal to that of the initial population (i.e., epistasis theorem [123]).

Still concerned with the convergence issue, a Markov chain analysis based on the

Schemata Theorem has been sketched in [57][153] where an infinite number of iterations is considered. Moreover, a convergence analysis with an infinite population size in also discussed in [159]. Furthermore, the effects of crossover have been thoroughly analyzed in [162].

Under genetic *drift* conditions § and in the case of a simple GA when mutation is either applied or not, the results in [110], numerically assessed in [3] through computer simulations, indicate that the mean convergence time grows proportionally to the population size. Concerning the mutation, an "optimal value" has been identified to allow all the solutions being explored with the same probability. Other empirical results about genetic drift for different versions of GAs can be found in [67][78].

Many other studies on the GA convergence and properties can be found in the state-of-the-art literature and are currently under development. The interested reader is referred to the specialized literature for a more complete discussion of these issues.

5.1.2. Differential Evolution - Since the main objective of DE is to improve the convergence rate of GAs, the main theoretical efforts have been addressed towards the optimal choice of the parameters controlling the evolution. This fact is confirmed by several works on this topic published in the reference literature (see [63][94][171][168][121][133] and references therein).

Since the basic idea of DE is to adapt the search step inherently along the evolutionary process to have a suitable trade-off between exploitation and exploration and the scale of the perturbation vectors is roughly proportional to the extent of the population diversity, the control parameters should allow large perturbations at the beginning of the evolution process when parental individuals are far away to each other. When the evolutionary process proceeds to the final stage, the population must be forced to a small region around the attraction basin of the global optimum through small perturbations. As a result, the adaptive search step would benefit the evolution algorithm by performing global search with a large perturbation step at the beginning of the optimization and refining the population with a small search step at the end.

In such a framework, although [151] states that the strategy parameters for the DE are not difficult to choose, there are not general rules for choosing the DE control coefficients. Moreover, even though there are only three parameters to set,

§ The random drift of the gene frequency is caused by the probabilistic generation of successive populations. It models the highlighting of genes with particular values.

the application of DE on several test functions as in [63] showed that finding the global optimum is very sensitive to the choice of the control variables: P (population size), ε (amplification factor), and p_C (crossover probability). Notwithstanding, the following rules of thumb have been given in [63]:

- a population size between $P_{min} = 3 \times N$ and $P_{max} = 8 \times N$;
- a good initial choice for the amplification factor $\varepsilon = 0.6$ to be increased if one suspects that this setting causes the trial solution being trapped in a local optimum. As a matter of fact, a larger ε increases the probability for escaping a local optimum, although for $\varepsilon > 1$ the convergence rate decreases since it is more difficult to reach the global solution when the perturbation is longer than the distance between two individuals;
- a large p_C often speeds-up convergence, but from a certain threshold value upwards the population may converge prematurely and stagnate. A good choice, whatever the cost function at hand, seems to be a value between 0.3 and 0.9.

Besides a careful analysis on the sensitivity of the DE optimization to the values of the control parameters, innovative operators have been also introduced by exploiting geometrical relationships to further speed up the convergence (e.g., trigonometric mutation [60]).

5.1.3. Particle Swarm Optimization In [42], Clerc and Kennedy examined in details the behavior of the PSO and defined some conditions on the PSO parameters to avoid a divergent search. With reference to a simplified one-dimensional (i.e., $N = 1$) and deterministic ($C_1 r_1 = C_1$ and $C_2 r_2 = C_2$) model, described by the following updating equations

$$\begin{aligned} v_{k+1} &= v_k + \varphi\,(t - f_k) \\ f_{k+1} &= f_k + v_{k+1} \end{aligned} \tag{43}$$

where $\varphi = C_1 + C_2$ and $t = \frac{C_1 p + C_2 g}{C_1 + C_2}$ is the index related to both the cognitive and the social term and by supposing the personal best and global best position fixed (i.e., $p_k = p$ and $g_k = g$), it has been shown that when $\varphi \geq 4$, the particles diverge as a function of k, while when $0 < \varphi < 4$ the trajectories are oscillating around the position t [112] with cyclic or quasi-cyclic behavior depending on φ. These conclusions have been

drawn from the analysis of (43) re-arranged in matrix form as follows: $\mathcal{F}_{k+1} = M\mathcal{F}_k$ where $\mathcal{F}_k = [v_k,\, z_k]^T$, being $z_k = (t - f_k)$, and the dynamic matrix is given by

$$M = \begin{bmatrix} 1 & \xi \\ -1 & 1-\xi \end{bmatrix}. \tag{44}$$

As a matter of fact, it turns out that $\mathcal{F}_k = M^k\mathcal{F}_0$, $\mathcal{F}_0$ being the initialization vector. A sufficient condition to reach an equilibrium point at the convergence (i.e., t) is that the amplitudes of the two eigenvalues of M are lower than unity [157]. However, a random choice of φ causes the uncontrolled increasing of the velocity term v_{k+1} [87].

Further developing the approach based on the generalized matrix, it has been proved that the following constriction system

$$\begin{aligned} v_{k+1} &= \chi\left[v_k + C_1 r_1\left(p - f_k\right) + C_2 r_2\left(g - f_k\right)\right] \\ f_{k+1} &= f_k + v_{k+1} \end{aligned} \tag{45}$$

where $\chi = \frac{2}{\left|2-\epsilon-\sqrt{\epsilon^2-4\epsilon}\right|} = 0.7298$ with $\varphi = 2C_1 = 2C_2 = 4.1$ guarantees the stability of the optimization process.

Other variants of the PSO exist and a careful analysis about the convergence taking into account the randomness of the algorithm has been reported in [120].

Concerning the optimal choice of the control coefficients, it is still worthwhile to point out that since higher values of ω produce relatively straight particle trajectories, resulting in a good global search characteristic, while small values of ω encourage a local searching, some researchers have gained advantage from a decrease [56][147] or a random variation of ω during the iterations [58]. As regards to the coefficients C_1 and C_2, they are usually set to 2.0 as recommended by some papers in the PSO literature [85][87][146] and found through experimentation in several optimization fields [15].

5.1.4. Ant Colony Optimization - A first proof on the convergence of an ACO-based algorithm, named graph-based ant system ($GBAS$), was reported in [72][73] where it has been shown that the global solution can be found at least once throughout the optimization process. Although reliable, such a proof does not hold true whatever the problem and it is limited to the $GBAS$ implementation which usually differs from the ACO version used in inverse scattering.

In [104], similarities between the pheromone update mechanism and the stochastic gradient descent have been pointed out to show that a class of *ACO* converges to a local optimum with probability equal to 1. On the same line of reasoning of [72], Stutzle and Dorigo proved in [152][54] that for a class of *ACO*-based algorithms with a lower bound ψ_{min} on the pheromone level, the success expectancy of the optimization is equal to $Pr\left\{\Phi_k^{opt} = \Phi^{best}\right\} \geq 1 - \eta$, with η small as desired and close to zero for an unlimited iterative process (i.e., $Pr\left\{\Phi_k^{opt} = \Phi^{best}\right\} \rightarrow 1$ when $k \rightarrow \infty$). Moreover, the pheromone deposited on the optimal path is higher than that left on others only after few iterations. A similar convergence proof has been also yielded in [172] by exploiting a simulated annealing concept and introducing an adaptive pheromone deposition function.

Besides the theoretical works on the convergence issues, some efforts have been also devoted to the application of the *ACO* to optimization problems not "suitable" for the structure of the algorithm itself. To describe this behavior, Blum and Dorigo [13] used the term "deception" previously introduced by Goldberg [68] to identify unfit problems for the *GA* concepts. The arising conclusions highlighted that in some cases *ACO* not only reaches a sub-optimal (local) solution (i.e., first order desception), but also that the performance of the algorithm can get worse (i.e., second order desception). For further indications on this issue, the interested reader is referred to the exhaustive survey on the *ACO* theory available in [55].

6. Evolutionary Algorithms - Applications to Inverse Scattering

In this section, the application of *EA*s to inverse scattering problems is analyzed. For notation simplicity and without loss of generality, the inverse scattering problems is formulated in two-dimensions and TM illuminations are considered to deal with a scalar system of equations. The extension to the vectorial $3D$ problem is straightforward and it does not modify the meaning and aim of the following discussion.

Since *EA*-based approaches have been applied to retrieve both dielectric and dissipative scatterers as well as perfectly conducting objects (PEC), the mathematical description of both problems as well as the analytical expression of the arising cost functions used during the optimization will be summarized. The theoretical reasons of the effectiveness of *EA*s in dealing with the ill-posedness and non-linearity of these problems will be discussed, as well.

The last part of this section provides a representative overview, to the best of

the authors' knowledge, on the solution of inverse scattering problems through *EA*s. Although a fair comparison among various algorithms and different implementations is impracticable due to (a) the customization of each *EA*s to the scattering scenario at hand, (b) different metrics adopted to define the cost function to be optimized, and (c) different strategies at the control level (e.g., different stopping criteria), a summary of the performance of some *EA*s is reported in Tabs. I and II. More specifically, Table I concerns with the *EA*-based approaches for qualitative imaging (i.e., the retrieval of the object's support and shape) of both conductors and dielectric scatterers. The values of the key computational indexes (where available) are given and compared: the number of unknowns, N, the number of trial solutions at each iteration, P, the number of iterations needed to achieve the convergence, K_{end}, and the corresponding total computational time, T_{tot}. Analogously, the performances of quantitative imaging (i.e., the retrieval of the dielectric properties within the investigation domain) techniques based on *EA*s are summarized in Tab. II.

6.1. Inverse Scattering of Dielectric Objects

Let us consider a region, called investigation domain D_i, characterized by a relative permittivity $\epsilon(\underline{r})$ and conductivity $\sigma(\underline{r})$. Such a region is probed by a set of V transverse-magnetic (TM) plane waves, with electric field $\underline{\zeta}^v(\underline{r}) = \zeta^v(\underline{r})\hat{\underline{z}}$ ($v = 1, \ldots, V$), $\underline{r} = (x, y)$, and the scattered field, $\underline{\xi}^v(\underline{r}) = \xi^v(\underline{r})\hat{\underline{z}}$, is collected at $M(v)$, $v = 1, ..., V$, measurement points $\underline{r}_{m(v)}$, $m(v) = 1, ..., M(v)$, distributed in the observation domain D_o.

In order to electromagnetically describe the investigation domain D_i, let us introduce the contrast function

$$\tau(\underline{r}) = [\epsilon(\underline{r}) - 1] - j\frac{\sigma(\underline{r})}{\omega\varepsilon_0} \quad , \quad \underline{r} \in D_i \tag{46}$$

where ω is the working angular frequency and the time dependence $e^{j\omega t}$ is supposed. Under the hypothesis of a linear, isotropic, and non-magnetic propagation medium, the scattered field $\xi^v(\underline{r})$ is the solution of the following Helmholtz equation (see [10])

$$\nabla^2\xi^v(\underline{r}) - \kappa^2(\underline{r})\xi^v(\underline{r}) = -j\omega\mu_0 J^v(\underline{r}) \tag{47}$$

where $\kappa(\underline{r}) = \omega\sqrt{\mu_o\epsilon_o\left[\tau(\underline{r}) + 1\right]}$ is the wavenumber. Moreover, $J^v(\underline{r})$ is the equivalent current density defined within D_i and radiating in free-space

$$J^v(\underline{r}) = \tau\left(\underline{r}\right) E^v\left(\underline{r}\right) \tag{48}$$

E^v being the electric field in the presence of the scatterer (i.e., the *total field*). By imposing that $\xi^v(\underline{r})$ satisfies the Sommerfeld's radiation condition

$$\lim_{r\to+\infty}\sqrt{r}\left(\frac{\partial\xi^v(\underline{r})}{\partial r}-j\kappa(\underline{r})\xi^v(\underline{r})\right)=0\,, \tag{49}$$

the solution of (47) in a two-dimensional scenario is given by the following Lippmann-Schwinger integral equations [35]

$$\xi^v\left(\underline{r}_{m(v)}\right)=\left(\tfrac{2\pi}{\lambda}\right)^2\int_{D_i}\tau\left(\underline{r}'\right)E^v\left(\underline{r}'\right)G_{2D}\left(\underline{r}_{m(v)}/\underline{r}'\right)d\underline{r}'\ ,\ \underline{r}_{m(v)}\in D_o\,, \tag{50}$$

$$\zeta^v\left(\underline{r}\right)=E^v\left(\underline{r}\right)-\left(\tfrac{2\pi}{\lambda}\right)^2\int_{D_i}\tau\left(\underline{r}'\right)E^v\left(\underline{r}'\right)G_{2D}\left(\underline{r}/\underline{r}'\right)d\underline{r}'\ ,\ \underline{r}\in D_i\,, \tag{51}$$

where λ is the background wavelength. Moreover, $G_{2D}\left(\underline{r}/\underline{r}'\right)$ is the two-dimensional free-space Green's function given by

$$G_{2D}\left(\underline{r}/\underline{r}'\right)=-\frac{j}{4}H_0^{(2)}\left(\frac{2\pi}{\lambda}\left\|\underline{r}-\underline{r}'\right\|\right)\,, \tag{52}$$

$H_0^{(2)}$ being the second-kind zeroth-order Hankel function.

Inverse scattering techniques are aimed at reconstructing the object function $\tau(\underline{r})$ in D_i starting from the knowledge of $\xi^v\left(\underline{r}_{m(v)}\right)$, $r_{m(v)}\in D_o$, and $\zeta^v\left(\underline{r}\right)$. Unfortunately, the arising problem is non-linear and ill-posed [10]. Moreover, a closed form solution of the integral equations in (50) and (51) does not generally exist. Consequently, the inverse scattering problem has to be reformulated and effective inversion methodologies have to be employed.

Since analytical solutions are rarely available, a numerical solution is then looked for. For instance, equations (50) and (51) are discretized according to the point-matching version of the Method of Moments [137]. The investigation domain D_i is partitioned into N square sub-domains D_n centered at $\underline{r}_n$, $n=1,...,N$. In each sub-domain, a pulse basis function is defined

$$\mathcal{B}_n\left(\underline{r}\right)=\begin{cases}1 & \text{if } \underline{r}\in D_n\\ 0 & \text{if } \underline{r}\notin D_n\end{cases}\ , \tag{53}$$

and the contrast function turns out to be expressed as follows

$$\tau\left(\underline{r}\right)=\sum_{n=1}^{N}\tau_n\mathcal{B}_n\left(\underline{r}\right)\ ,\quad \underline{r},\,\underline{r}_n\in D_i \tag{54}$$

where $\tau_n=\tau\left(\underline{r}_n\right)$, $n=1,...,N$. By assuming the incident field ζ^v and the total field E^v constant inside each sub-domain D_n, the discrete form of the Lippmann-Schwinger equations is given by

$$\xi^v_{m(v)}\left(\underline{r}_{m(v)}\right)=\textstyle\sum_{n=1}^{N}\tau_nE^v_n\left(\underline{r}_n\right)G_{2D}\left(\underline{r}_{m(v)}/\underline{r}_n\right)\ ,\ \underline{r}_{m(v)}\in D_o\,, \tag{55}$$

$$\zeta_n^v\left(\underline{r}_n\right) = E_n^v\left(\underline{r}_n\right) - \sum_{p=1}^{N} \tau_p E_p^v\left(\underline{r}_p\right) G_{2D}\left(\underline{r}_n/\underline{r}_p\right) \quad , \quad \underline{r}_n \in D_i \,, \tag{56}$$

where $G_{2D}\left(\underline{r}_m/\underline{r}_n\right)$ is the discretized form of the two-dimensional Green's operator.

In order to cope with ill-posedness, the inverse scattering problem is usually recast as an optimization one defining a suitable cost function proportional to the mismatch between the measured fields and their numerically evaluated counterparts to be minimized. Additional regularization or penalty terms can be also added to the cost function in order to enhance the reliability of the inversion process. The cost functional is a function of the trial solution $\underline{f} = \left\{\widetilde{\tau}_n,\, \widetilde{E}_n^v;\, n = 1, ..., N\right\}$ and it can be expressed in matrix form as follows [10]

$$\Phi\left\{\underline{f}\right\} = \alpha \frac{\sum_{v=1}^{V} \left\| \underline{\xi}^v - \underline{\underline{G}}_{EXT}^v \widetilde{\underline{\tau}}\, \widetilde{\underline{E}}^v \right\|^2}{\sum_{v=1}^{V} \left\| \underline{\xi}^v \right\|^2} + \beta \frac{\sum_{v=1}^{V} \left\| \underline{\zeta}^v - \underline{\widetilde{E}}^v + \underline{\underline{G}}_{INT}^v \widetilde{\underline{\tau}}\, \widetilde{\underline{E}}^v \right\|^2}{\sum_{v=1}^{V} \left\| \underline{\zeta}^v \right\|^2} \tag{57}$$

where $\underline{\underline{G}}_{EXT}^v$ and $\underline{\underline{G}}_{INT}^v$ are the $M \times N$ external Green's matrix and the $N \times N$ internal Green's matrix, respectively. Moreover, α and β are two user-defined regularization parameters. Furthermore, $\underline{\zeta}^v$ is the $N \times 1$ incident field array, the $M \times 1$ entries of $\underline{\xi}^v$ are given by the measured scattered field samples, and $\underline{\widetilde{E}}^v$ is the $N \times 1$ array of the estimated total electric field.

The actual solution $\underline{f}^{opt}$ is looked for as the $N \times 1$ trial array that minimizes the cost function (57)

$$\underline{f}^{opt} = \arg\left\{\min_{k=1,...,K}\left[\Phi\left\{\underline{f}_k\right\}\right]\right\} \tag{58}$$

where $\underline{f}_k = \left\{\widetilde{\underline{\tau}}_k,\, \underline{\widetilde{E}}_k^v\right\}$ is the trial solution at the step k-th iteration of the optimization procedure.

6.2. Inverse Scattering of Perfect Electric Conductors

When dealing with *PEC* characterized by a conductivity $\sigma \to \infty$, Equation (50) modifies as follows

$$\xi^v\left(\underline{r}_{m(v)}\right) = -\tfrac{\omega\mu_0}{4} \oint_\gamma J_s^v\left(\underline{r}'\right) G_{2D}\left(\underline{r}_{m(v)}/\underline{r}'\right) d\underline{r}' \quad , \quad \underline{r}_{m(v)} \in D_o \,, \tag{59}$$

$J_s^v\left(\underline{r}\right)$ being the surface current density defined only on the boundary γ of the unknown scatterer. Since the following condition holds true on the surface of the *PEC*

$$\xi^v\left(\underline{r}\right) + \zeta^v\left(\underline{r}\right) = 0 \quad , \quad \underline{r} \in \gamma, \tag{60}$$

the scattering equation is given by

$$\zeta^v\left(\underline{r}\right) = \tfrac{\omega\mu_0}{4} \oint_\gamma J_s^v\left(\underline{r}'\right) G_{2D}\left(\underline{r}/\underline{r}'\right) d\underline{r}' \quad , \quad \underline{r} \in \gamma \,, \tag{61}$$

and the unknown current $J_s^v(\underline{r})$ descriptive of the scatterer (i.e., the contour γ) is computed through the inversion of the linear system (61) starting from the knowledge of $\zeta^v(\underline{r})$. Likewise the inversion of dielectric scatterers, the reconstruction process is recast as the minimization of the following cost function

$$\Phi\left\{\underline{\gamma}\right\} = \frac{\sum_{v=1}^{V}\left\|\underline{\xi}^v - \underline{\underline{G}}_{EXT}^v \tilde{\underline{J}}_s^v(\underline{\gamma})\right\|^2}{\sum_{v=1}^{V}\left\|\underline{\xi}^v\right\|^2} \tag{62}$$

where $\underline{\gamma}$ is the parametrized representation of the scatterer contour γ.

Because of the non-linearity of the scattering problem and the presence of local minima (i.e., false solutions of the inverse scattering problem) in the cost function (57) or (62), the quality and the reliability of the final solution mostly depends on the effectiveness of the search strategy.

6.3. EAs-based Approaches for Inverse Scattering

The first multiple-agent evolutionary techniques applied to solve microwave inverse scattering problems were the genetic algorithms. Chiu and Liu in [39] applied the BGA for the $2D$ inversion of a PEC cylinder illuminated by an incident TM-polarized plane wave. The $2D$ surface-reconstruction problem has been reformulated into a mono-dimensional one by describing the contour of the cylinder as a function of the polar angle θ

$$\gamma(\theta) = \sum_{m=0}^{M/2} A_m cos(m\theta) + \sum_{m=1}^{M/2} B_m sin(m\theta) \tag{63}$$

with $\theta \in [0, 2\pi]$, where the unknowns to be determined

$$\underline{f} = \left\{A_0,\, A_1,\, A_{\frac{M}{2}},\, B_1,\, B_2, ...,\, B_{\frac{M}{2}}\right\} \tag{64}$$

are the real coefficients of the Fourier series expansion. The number of unknown parameters was set to $N = M + 1 = 9$ and various experiments considering strings of length $l = 8 \times N$ and $l = 10 \times N$ have been performed to validate the EA-based inversion method. As far as GA parameters are concerned, a population of $P = 300$ individuals was chosen with $p_C = 0.8$ and $p_M = 0.04$. Successively, the sensitivity of the reconstruction on the GA parameters has been analyzed in [40] by the same authors under TE illuminations. The outcomes have been that for this kind of problems a suitable choice of the GA parameters was: a population dimension in the range $P \in [300,\, 600]$, a chromosome length of $l \in [8,\, 16] \times N$ bits, and probabilities of crossover

and mutation in the following ranges $0.7 < p_C < 0.9$ and $5 \times 10^{-4} < p_M < 5 \times 10^{-2}$, respectively. Such a GA-based inversion procedure was extended in [41] to image lossy or imperfectly conducting cylinders. More specifically, the GA was used to retrieve also the conductivity of the unknown scatterer by coding both the Fourier coefficients of the shape and the value of the conductivity of the object.

Following the guidelines in [39] for defining the inversion problem, an approach based on a micro-GA has been presented in [79] to enhance reconstruction and convergence performances of standard GAs. A number of $N = M + 1 = 5$ unknowns was considered to describe the contour of the scatterer through (63) and a population of $P = 5$ individuals with $l = 12 \times N$ chromosomes was used. The main advantages of the μGA are that it employs a small population, thus reducing the overall computational burden, and assures a fast convergence to sub-optimal solutions while maintaining superior search ability.

Takenaka and co-workers [154] proposed a volume-reconstruction approach to estimate widths and locations of parallel strips in $1D$ and $2D$ problems without any a-priori information on the number of strips. The investigation domain D_i was discretized in $N = 20, 36$ cells and either an empty ("$a_i = 0$") or occupied ("$a_i = 1$") state was assigned to each cell. If $a_i = 1$, the ith cell is occupied by a metallic strip, $a_i = 0$ otherwise. It is worth pointing out that, the original problem was reformulated as the definition of a binary map to allow a straightforward use of the BGA. The population dimension was set to $P = 50$ with fixed crossover probability $p_C = 0.8$ and variable mutation probability in the range $p_M \in [0.01, 0.5]$. The approach was extended in [100] to retrieve locations and two-dimensional cross sections of conducting cylinders. To improve the convergence rate for a $2D$ discretization of the scenario under test, a customized crossover, called *rectangular block* crossover, was developed to efficiently deal with a binary reconstruction map. Moreover, a larger population with $P = 200$ individuals was considered because of the wider solution space ($N = 225$). In [173], a similar BGA-based approach has been tested against Ipswich experimental data-set to reconstruct metallic objects. The investigation domain D_i was discretized into $N = 400$ cells and the GA parameters were set to $P = 100$, $p_C = 0.8$, and $p_M = 0.2$.

To avoid the quantization error related to the discretization of the real coefficients in (63), Qing and co-workers proposed in [124][125][127] a strategy based on a RGA. For comparative purposes, some benchmark examples previously addressed in [39] with the BGA have been considered. An experimental validation of the method has been

presented in [128], as well.

In an alternative fashion, the contour of the conducting cylinders has been approximated in [129] by means of local shape functions mathematically expressed in terms of closed cubic B-splines

$$\gamma(\theta) = \sum_{m=0}^{M-1} \rho_m \left(\frac{M}{2\pi}\theta - m \right), \quad \theta \in [0, 2\pi] \tag{65}$$

where each segment ρ_m is a linear combination of four cubic polynomials $Q_i(t)$, $i = 0, ..., 3$, as follows

$$\rho_m(t) = p_{m-1}Q_0(t) + p_m Q_1(t) + p_{m+1}Q_2(t) + p_{m+2}Q_3(t)$$

$p_{m-1}, ..., p_{m+2}$ being the control points. In this case, the parameters to be optimized are the set of control points

$$\underline{f} = \{p_0, ..., p_{M-1}\} \ . \tag{66}$$

Dealing with these problems, the RGA was used with probability coefficients equal to $p_C = 1.0$ and $p_M = 0.1$. More recently, the representation of PEC contours using cubic splines has been extended to deal with three-dimensional ($3D$) electrically large conducting patches [141].

Previous references and the obtained results point out that the GA-based techniques have demonstrated to work effectively in retrieving strong scatterers in free-space through the minimization of the mismatch between the measured and the reconstructed scattered field. In addition, the robustness of the GA in such a framework has been proved since the converge to the global optimum has been obtained with high probability despite a rough initialization of the iterative process. More recently, an innovative strategy based on Genetic Programming (GP) [88] has been presented in [163]. A new geometry-encoding scheme was introduced and a tree-shaped chromosome was used to describe the shapes of the cylinders as the union and subtraction of convex polygons.

In [65], the binary GA was also applied to both the detection of circular conducting cylinders buried in an homogeneous dielectric medium and the dielectric profile retrieval of layered media. A similar approach has been considered in [101] to reconstruct the electrical parameters (ϵ_i, σ_i, μ_i, $i = 1, ..., M$) of a multilayered radome of finite size Δ. In this case, the chromosome was the binary representation of the following unknown array

$$\underline{f} = \{\epsilon_1, \sigma_1, \mu_1, d_1, ..., \epsilon_M, \sigma_M, \mu_M, d_M\} \tag{67}$$

being $\sum_{n=1}^{N} d_n = \Delta$. To increase the accuracy of the reconstruction and improve the convergence rate, an adaptive chromosome structure was chosen in order to iteratively adjust the existence range of each parameter.

An inverse scattering technique for the detection of perfectly conducting cylindrical objects buried in a half-space have been described in [38]. An improved "steady-state" GA $(SS-GA)$ ‖ was used to reduce the computational burden and a non-uniform probability distribution was introduced to control the generation of new individuals through crossover and mutation. The shapes of the buried objects have been represented considering both Fourier series (63) and Cubic-splines (65) representations. Each unknown was coded with $l = 20$ bits string and a population of $P = 100$ individuals was chosen. Moreover, the following setup was chosen: $p_C = 0.05$ and $p_M = 0.5$. If compared to standard GA, the values of the control parameters, turn out different in magnitude. This is due to the steady-state GA implementation since only a portion of new individuals is generated through crossover and mutation, while the whole population is updated in standard GA. In [38], it has also been verified that even for an initial guess far away from the optimal solution the computational cost to reach the global solution is much less in $SS-GA$ than for simple GAs. Moreover, a further reduction of the computational time was obtained running the optimization process in parallel on a multiprocessor cluster system.

In the framework of subsurface imaging, a GA-based approach for the retrieval of the dimension and location of a $3D$ buried object has been presented in [93]. A parallel binary GA procedure has been considered to speed up the fitness evaluation computed through the $FDTD$ (Finite-Difference Time-Domain) method. More in detail, the cost function has been defined as the difference between the measured and the calculated s_{11} parameters on a frequency band from ν_L up to ν_H at the port of the probing antenna

$$\hat{\Phi} = 1 - \frac{\sqrt{\sum_{\nu=\nu_L}^{\nu_H} \left[s_{11}^{meas}(\nu) - s_{11}^{calc}(\nu) \right]^2}}{\sqrt{\sum_{\nu=\nu_L}^{\nu_H} \left[s_{11}^{meas}(\nu) \right]^2}}. \tag{68}$$

As far as the GA is concerned, a population of $P = 50$ individuals and $p_C = 0.5$ and $p_M = 0.2$ were chosen.

Besides shape reconstruction problems, the reconstruction of the dielectric properties of unknown objects has been faced in [23] by means of a quantitative (pixel

‖ In steady-state GAs, only a portion of the population is updated and a suitable replacement strategy is considered.

based) microwave imaging technique. A binary GA with $Q = 256$ quantization levels has been used to describe the real-valued unknowns. The inverse scattering problem has been solved in the framework of the Born approximation to reduce both the non-linearity of the descriptive scattering equations as well as the total number of unknowns. The chromosome to be optimized was

$$\underline{f} = \{\tau_1, \tau_2, ..., \tau_n, ..., \tau_N\} \tag{69}$$

where $N = 900$. The simulations were carried out with a population of $P = 100$ individuals and crossover and mutation probability coefficients were set to $p_C = 0.7$ and $p_M = 4 \times 10^{-4}$, respectively. Moreover, a sensitivity analysis was performed varying the control parameters in the following range: $P \in [40, 200]$, $0.6 < p_C < 0.8$, and $4 \times 10^{-4} < p_M < 10^{-3}$. It has been proved that for small-sized populations the quantitative errors increase as for either low value of p_C or high p_M. The same optimization procedure has been validated in [115] against experimental data acquired when considering highly-contrasted bodies. Since, the Born approximation cannot be applied, the unknown vector was composed of both the contrast function $\underline{\tau}$ and the total field $\underline{E}^v$ in the investigation domain

$$\underline{f} = \{\tau_1, \tau_2, ..., \tau_N, E_1^v, E_2^v, ..., E_N^v; v = 1, ..., V\}. \tag{70}$$

Because of the continuous nature of the parameters to be optimized, a real-coded GA has been proposed in [24] and significantly superior performances with respect to the BGA have been attained [25]. The potentialities of the RGA have been further pointed out and the methodology extended to hybrid-coded chromosomes in order to deal with both nondestructive testing and evaluation ($NDT-NDE$) problems [26] and biomedical imaging [29]. More specifically, the unknowns were expressed by means of the following vector

$$\underline{f} = \{x_0, y_0, L, W, \theta, E_1^v, E_2^v, ..., E_M^v; v = 1, ..., V\} \tag{71}$$

coding through binary strings the values of the barycenter (x_0, y_0), the length (L), the width (W), and the orientation (θ) of either a crack in $NDT - NDE$ problems or a pathology in case of biomedical imaging. Differently, a floating-point representation has been used to code the unknown field values. Reliable values of the probability coefficients for the RGA turned out to be $p_C = 0.7$ and $p_M = 0.4$.

In $NDT - NDE$ problems, the application of GAs starts with a BGA proposed in [2] to identify the qualitative nature (i.e., length, width, orientation, and barycenter)

of a crack on the surface of an object. In order to tackle more complex diagnosis problems, an innovative description of the crack based on a suitable parameter selection (71) as well as a more effective exploitation of the a-priori information has been considered in [116][30][7] to reduce the number of problem unknowns and enable an efficient use of HGAs. Although effective, these approaches considered scattering configurations characterized by the presence of only a single defect. To overcome such a limitation, two enhanced GA-based optimization techniques able to deal with multiple defects in a dielectric host medium have been proposed in [8]. Both methods adopted a multicrack variable-length hybrid coding. The former strategy was based on a hierarchical implementation, which considers a set of parallel sub-processes, each one looking for a solution with a fixed number of cracks. The other deals with a single optimization process aimed at retrieving the best reconstruction among different crack-length solutions. Because of the use of an ad-hoc operator to correctly recombine the discrete (binary) and continuous part of the chromosome, the control probabilities were kept constant to $p_C = 0.7$ and $p_M = 0.4$ for each portion of the chromosome structure. Unlike [8], also the reconstruction of the dielectric properties of the defects has been addressed in [9].

Similar concepts have been exploited to deal with biomedical imaging problems as discussed in [29]. The chromosome structure was still chosen as a two-part variable-length string. In such a case, the variable-length structure was used because of the variable number of discretization sub-domains occupied by the pathology where the unknown field has to be computed.

A parallel $SS - GA$ integrated with a $FDTD$ approach has been presented in [167][140] for early cancer detection. Parallel computing was considered due to the large computational burden of the $FDTD$-based approach caused by the fine discretization of the investigation domain ($N = 600 \times 600$ cells).

As regards to the efforts devoted to increase the computational efficiency of GA-based inversions, ad-hoc versions or specific operators have been designed. Representative examples of a wide literature are the use of nonuniform probability densities in BGAs [37] and a parabolic crossover operator for RGAs [18].

Despite the success of GAs-based approaches in several area of electromagnetic and inverse scattering, more recently other EAs have shown better performance. Due to its faster convergence with respect to GAs, the DE was alternatively used to face electromagnetic inversion and it has been firstly applied to image circular-cylindrical

conductors and tunnels [105][106]. Later, the DE was used to solve benchmark problems [130] and its performances were compared to the solutions from the RGA in [127]. A population of $P = 5 \times N$ individuals was considered. Moreover, the crossover probability and the mutation intensity were set to $p_C = 0.9$ and $\epsilon = 0.7$, respectively. Without *a priori* information on the number of cylinders within the investigation region, the DE strategy with individuals in groups ($GDES$) has been proposed [131]. The key idea of the $GDES$ is to organize the population into different groups. The individuals of the same group code the same number of cylinders and have the same chromosome length. Successively, an innovative DE-based algorithm was proposed by Qing [132]. In the dynamic DE strategy ($DDES$), a larger (virtual) population has been chosen to speed up the convergence. The new individuals generated at iteration $k+1$ compete during the same iteration with their parents. As a consequence, the mating operation turns out being more sensitive to the fast changes of the population with a consequent enhancement of the convergence rate.

Because of the strong dependence of the DE performance on both its control parameters and the cost function to be optimized, an extensive calibration of the population size, P, the crossover probability, p_C, and the mutation intensity, ϵ, has been carried out in [133] specifically for imaging problems concerned with PEC cylinders in free space. More recently, a comparative study on the efficiency of DE and PSO when applied to the shape reconstruction of PEC scatterers has been reported in [134].

In [97], an approach for the detection of $2D$ buried inhomogeneities has been designed by combining two DE-based optimization techniques. More specifically, the DE/1/best/bin version ($p_C = 0.8$ and $\epsilon = 0.6$) was used to rapidly locate the attraction basin of a minimum and successively the algorithm switched to the DE/1/rand/bin ($p_C = 1.0$ and $\epsilon = 0.6$) to avoid the trial solution be trapped in a local minimum.

DE has been also applied to the $3D$ detection of unexploded ordnance (UXO) [34] and lossy spherical objects buried in the subsoil [6]. In this latter case, a modified DE algorithm was considered where multiple populations evolve in parallel analogously to [131].

More recently, EAs inspired by the foraging behavior of swarms have proved to outperform previous EAs in dealing with a set of imaging problems also related to high-dimensional continuous spaces.

A standard PSO algorithm has been used in [145] to reconstruct $1D$ permittivity and conductivity profiles in lossy and inhomogeneous media. Because of its ability in

exploring the parameter space and avoiding wrong solutions thanks to the cooperative behavior of the agents, the PSO has been also profitably considered for the reconstruction of $2D$ dielectric scatterers [31][49][50][62]. The enhanced convergence rate of the PSO with respect to the RGA has been assessed both in [49] and [50]. Moreover, the calibration of the PSO control parameters carried out in [49] has proved that the most suitable setup (in terms of decrease of the cost function in a fixed amount of iterations) is: $\omega = 0.4$ and $C_1 = C_2 = 2.0$, in accordance with the outcomes of other published papers (see Section 5.1.3). Moreover, the ratio $\frac{N}{P} \cong 5.5$ has been deduced as a good rule of thumb for the size P of the swarm for this class of optimization problems.

The retrieval of $3D$ lossy dielectric objects has been addressed in [80][81]. Due to the limited set of independent scattering data and the dimensionality of the problem at hand, an adaptive multiresolution technique was integrated into the swarm evolution to reduce the search space and make more efficient the PSO-based minimization [51]. In order to deal with high-dimensional spaces, an alternative approach based on a μPSO, employing a reduced swarm in analogy with the μGA, has been also presented in [83]. Otherwise, to avoid the premature convergence of the standard version of the PSO, modified social structures have been envisaged [82]. More specifically, besides the standard version of the PSO (i.e., that presented in this work), whose structure is known as *gbest* topology since the information are instantaneously communicated to the whole swarm, in [82] other topologies are considered in order to limit the communication between the particles [99] to prevent premature convergence.

The PSO has been also successfully applied in industrial and biomedical imaging problems. For example, a swarm-based reconstruction algorithm for the detection and characterization of multiple inclusions in concrete structures has been presented in [156]. Moreover, a PSO-based technique for early cancer detection has been discussed in [170].

In the framework of swarm based approaches, only preliminary results are available (e.g., [118]) concerning the application of the ACO to inverse scattering even though an interesting hybridization with the linear sampling method (LSM) has been recently studied in [20] to inspect $3D$ homogeneous dielectric scatterers.

As far as hybrid algorithms are concerned, a μGA and a RGA coupled with a local search method have been presented in [166] and in [126], respectively, for imaging perfectly conducting cylinders. In [174], the GA was combined with a tabu mechanism to avoid a (repetitive) sampling of poor regions within the search space. Following the same guidelines outlined in [163], the authors combined a GA-inspired optimization

with a local search method [164] reducing the number of cost function evaluations to reconstruct bowtie-shaped conducting cylinders from 1.17×10^5 [163] down to 5421. An hybrid-GA has been also proposed in [66] to retrieve the dielectric profile of a layered medium.

In the framework of quantitative imaging problems, the permittivity reconstruction of $2D$ objects with large size and high contrast has been carried out in [114] by combining a Levenberg-Marquardt algorithm with the GA. To speed-up the convergence, a Polak-Ribière conjugate gradient (CG) has been merged into a global optimization loop performed with a RGA in [24][116][117]. Furthermore, a parallel implementation of such a hybrid technique has been detailed in [98].

As regards to memetic algorithms, they have been used to detect cylindrical inhomogeneities starting from phaseless data obtained by synthetic as well as experimental measurements [28]. Moreover, a MA-based approach has shown to effectively work for the electromagnetic reconstruction of buried objects [27][32], as well. Because the heavy computational burden, MAs applications are usually limited to low-dimensionality problems.

7. Summary and Conclusions

In this paper, a review of Evolutionary Algorithms as applied to inverse scattering problems has been reported. After an introduction on the genesis of EAs (Sect. 2), the most representative and widespread evolution-based techniques have been described in a common framework and by means of a uniform notation (Sect. 3) to point out the main similarities and differences among the various implementations detailed in Sect. 4. Some theoretical hints concerning the convergence properties and the parameters selection have been also discussed (Sect. 5). Section 6 has been devoted to present state-of-the-art applications of EAs to electromagnetic imaging problems. Such a critical discussion has pointed out that the success of EAs in dealing with nonlinear ill-posed inverse scattering problems mainly relies in a suitable set of answers to the following key-issues:

- a suitable representation of the unknowns and a proper choice of the EA is mandatory starting from a careful analysis of the problem at hand and its numerical description (i.e., the cost function). According to the "No free lunch theorem" [165], the "optimum" algorithm does not exist since the average performance of any pair

of algorithms across all the possible problems is identical;

- physical constraints need to be taken into account to enhance the effectiveness of *EA*s by reducing the area of the solution space to be sampled during the optimization;
- the *a-priori* knowledge on the scenario under test needs to be profitably incorporated into both the solution representation and the evolutionary operators to guide the search process and increase its convergence rate;
- great care must be exercised in defining the cost function since it represents the only link between the physical problem and its numerical counterpart. Failing such a definition prevents the actual solution is reached at the convergence of the *EA* to its global optimum;
- the calibration of the evolutionary procedures needs to be carefully performed to fully exploit the *EA* potential. On the other hand, it should be stressed that no single-test-case calibration is necessary, but the tuning of the control parameters must be carried out on a class of problems (e.g., imaging of dielectric objects) to avoid "overfitting" and confer generalization features on the *EA*;
- the feasibility and reliability of an *EA* must be assessed first on a benchmark of test functions and then in comparison with other deterministic and stochastic microwave imaging techniques.

8. Open Problems and New Research Developments

As far as the application of *EA*-based microwave imaging techniques to inverse scattering problems is concerned, it should be first pointed out that the development of evolutionary techniques has received a great boost in the last twenty years due the continuous enhancement of the computational capabilities of modern personal computers, but also for their flexibility and features usually very suitable to face with the ill-posedness and nonlinearity of the arising optimization problem. As a matter of fact,

- *EA*s are multiple-agent optimizers;
- *EA*s are global and hill-climbing algorithms thanks to their stochastic nature;
- *EA*s allow the straightforward introduction of *a-priori* information or constraints;

- *EA*s are able to deal with floating-point and/or discrete and/or binary unknowns simultaneously;
- *EA*s are intrinsically parallel algorithms;
- *EA*s are easily integrated with local optimizers.

However, some other drawbacks limit their effectiveness besides typical negative issues of inverse scattering problems. For example,

- the computational burden (especially when moving towards $3D$ scenarios);
- the low convergence rate when close to the global solution although in its attraction basin;
- the dependence on the parametrization of the problem unknowns;
- the sensitivity to the calibration parameters.

As regards to the computational issues, some receipts to limit these drawbacks consist in: (a) reducing the number of problem unknowns by recurring to a suitable parametrization of the scatterer under test [26][30] or considering a multi-resolution strategy [50][51] or a multi-stage reconstruction [108]; (b) hybridizing the EA with a deterministic optimizer [114][24][98]; (c) computing at each iteration the secondary unknowns (i.e., the field distribution within the investigation domain) by means of fast forward solvers (see [36] and the reference therein); (d) exploiting the explicit parallelism of *EA*s through a parallel implementation [98].

With reference to the EA parallelization (d), which has been left out in the main body of the paper, it is well known that one of the most attractive features of nature-inspired optimization techniques is their parallelism that allows an effective sampling of the solution space. Besides the implicit parallelism still exploited in serial implementations, the parallelism of an EA is also guaranteed by its multiple-agent nature. As a matter of fact, a number of sample points equal to the population dimension is processed at each iteration to effectively look for the global optimum. In order to fully exploit also this characteristic, a parallel implementation of the iterative procedure would fully exploit also this characteristic enabling (i) a parallel and simultaneous search from multiple points in the solution space; (ii) a more efficient search, even when no parallel hardware is available; (iii) a higher efficiency than sequential implementation, and (iv) a speedup due to the use of multiple CPUs.

Despite these envisaged advances, it should be also pointed out that the use of a parallelized (bare/hybrid) EA is not different from other parallel methodologies and its efficiency largely depends upon the system architecture, the parallel execution overhead, the number of new agents created at each iteration, the population structure, and the parallel *granularity* (i.e., the CPU time of the steps being executed in parallel). These advances can be reached only if:

- a *structured population* [21] is taken into account to obtain not only a faster algorithm, but also a superior numerical optimization able to profitably exploit the *multi-agent* nature of the EA;
- some agents do a different local search (*decentralized local optimization*) in order to improve the convergence rate of the iterative process;
- the evolution process explicitly keeps memory of the population evolution in order to reduce/avoid the runtime of the cost function evaluation for similar/equal individuals;
- the evolutionary operators are applied in parallel.

As regards to the enhancement of the convergence rate through the reduction of the extension of the solution space to be sampled during the optimization, the number of iteration K_{end} evidently reduces in correspondence with an increase/efficient-exploitation of the *a-priori* information. Indeed, an additional information on the location of the attraction basin of the global solution usually helps the evolutionary procedure in locating the actual solution as well as the EA designer in defining the optimal parametrization of the problem unknowns.

Another way to save computational resources when applying EAs to inverse scattering problems is to use a succession of inversion procedures, each one concerned with a number of unknowns smaller or equal than the information content of the scattering data in order to "simplify" the cost function to be optimized. The reduction of the complexity of the cost function can be yielded in different ways according to some recently developed strategies. In such a framework, it is worthwhile to mention multi-resolution methods [50][51] devoted to perform an iterative synthetic zoom over the region where the scatterer is supposed to be located and multi-stage reconstructions [108][20] where each inversion is aimed at identifying different characteristics of the unknown scatterer until its complete description/knowledge.

Nevertheless, although further theoretical and numerical developments are still required to consider EA-based inversion techniques mature, reliable, and efficient inversion methodologies, the expected impact of such approaches justifies significant research effort to develop EA-based tools dedicated to specific applications in the framework of microwave imaging and NDT/NDE applications.

Acknowledgements

The authors thank very much B. Symes and and *Inverse Problems* for the kind invitation to write this review paper and Ch. Benson and K. Watt for their cooperation and assistance. The authors are indebted to D. Lesselier and they would like to acknowledge useful discussions on inverse scattering problems and optimization techniques with T. Takenaka, Ch. Pichot, A. Abubakar, O. Dorn, M. El-Shenawee, K. Belkebir, M. Saillard, I. Akduman, P. Meaney, J. Ch. Bolomey, R. Zoughi, M. Pastorino, and T. Isernia. Finally, A. Massa would like to express his everlasting gratitude to E. Vico.

References

[1] M. M. Ali and A. Torn, "Population set based global optimization algorithms: Some modifications and numerical studies," *Computers and Operations Research*, vol. 31, 1703-1725, 2004.

[2] A. A. Arkadan, T. Sareen, and S. Subramaniam, "Genetic algorithms for nondestructive testing in crack identification," *IEEE Trans. Mag.*, vol. 30, pp. 4320-4322, 1994.

[3] H. Asoh and H. Muhlenbein, "On the mean convergence time of evolutionary algorithms without selection," in *Parallel Problem Solving from Nature III*, vol. 866, pp. 98-107, 1994.

[4] T. Back, "Selective pressure in evolutionary algorithms: A characterization of selection mechanisms," in *Proc. 1st IEEE Conf. on Evolutionary Computation*. Piscataway, NJ: IEEE Press, 1994, pp. 57-62.

[5] A. L. Barker and W. N. Martin, "Dynamics of a distance-based population diversity measure," in *Proc. 2000 IEEE Congress on Evolutionary Computation*, 16-19 Jul. 2000, vol. 2, pp. 1002-1009.

[6] A. Bréard, G. Perrusson, and D. Lesselier, "Hybrid differential evolution and retrieval of buried spheres in subsoil," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, pp. 788-792, 2008.

[7] M. Benedetti, M. Donelli, G. Franceschini, M. Pastorino, and A. Massa, "Effective exploitation of the a-priori information through a microwave imaging procedure based on the SMW for NDE/NDT applications," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, pp. 2584-2592, 2005.

[8] M. Benedetti, M. Donelli, and A. Massa, "Multicrack detection in two-dimensional structures by means of GA-based strategies," *IEEE Trans. Antennas Propag.*, vol. 55, pp. 205-215, 2007.

[9] M. Benedetti, G. Franceschini, R. Azaro, and A. Massa, "A numerical assessment of the reconstruction effectiveness of the integrated GA-based multicrack strategy," *IEEE Antennas Wireless Propag. Lett.*, vol. 6, pp. 271-274, 2007.

[10] M. Bertero and P. Boccacci, *Introduction to Inverse Problems in Imaging*. Bristol, U.K.: IOP Press, 1998.

[11] A. Bertoni and M. Dorigo, "Implicit parallelism in genetic algorithms," *Artif. Intell.*, vol. 61, pp. 307-314, 1993.

[12] B. Bilchev and I. C. Parmee, "The ant colony metaphor for searching continuous design spaces," in *Proc. AISB Workshop Evol. Comput.*, vol. 993 of Lecture Notes in Computer Science, pp. 25-39, 1995.

[13] C. Blum and M. Dorigo, "Search bias in ant colony optimization: on the role of competition-balanced systems," *IEEE Trans. Evol. Comput.*, vol. 9, pp. 159-174, 2005.

[14] C. Blum, M. J. Blesa Aguilera, A. Roli, and M. Sampels, *Hybrid Metaheuristics: An Emerging Approach to Optimization*. Berlin Heidelberg: Springer, 2008.

[15] D. W. Boeringer and D. H. Werner, "Particle swarm optimization versus genetic algorithms for phased array synthesis," *IEEE Trans. Antennas Propag.*, vol. 52, pp. 771-779, 2004.

[16] E. Bonabeau, M. Dorigo, and G. Theraulaz, *Swarm Intelligence: From Natural to Artificial Systems*. New York, NY: Oxford Univesity Press, 1999.

[17] L. B. Booker, D. E. Goldberg, and J. H. Holland, "Classifier systems and genetic algorithms," *Artif. Intell.*, vol. 40, pp. 235-282, 1989.

[18] E. Bort, G. Franceschini, A. Massa, and P. Rocca, "Improving the effectiveness of GA-based approaches to microwave imaging through an innovative parabolic crossover," *IEEE Antennas*

Wireless Propag. Lett., vol. 4, pp. 138-142, 2005.

[19] M. F. Bramlette and E. E. Bouchard, "Genetic algorithms in parametric design of aircraft," in *Handbook of Genetic Algorithms*. New York: Van Nostrand Reinhold, 1991, ch. 10, pp. 109-123.

[20] M. Brignone, G. Bozza, A. Randazzo, M. Piana, and M. Pastorino, "A hybrid approach to 3D microwave imaging by using linear sampling and ACO," *IEEE Trans. Antennas Propag.*, vol. 56, pp. 3224-3232, 2008.

[21] E. Cantu-Paz, "A summary of research on parallel genetic algorithms," *Tech. Rep. 95007*, IlliGal Rep., 1995.

[22] S. Caorsi, A. Massa, M. Pastorino, and F. Righini, "Crack detection in lossy two-dimensional structures by means of a microwave imaging approach," *Int. Journ. Applied Electromagnetics Mechanics*, vol. 11, no. 4, pp. 233-244, 2000.

[23] S. Caorsi and M. Pastorino, "Two-dimensional microwave imaging approach based on a genetic algorithm," *IEEE Trans. Antennas Propag.*, vol. 48, pp. 370-373, 2000.

[24] S. Caorsi, A. Massa, and M. Pastorino, "A computational technique based on a real-coded genetic algorithm for microwave imaging purposes ," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, pp. 1697-1708, 2000.

[25] S. Caorsi, A. Costa, and M. Pastorino, "Microwave imaging within the second-order born approximation: stochastic optimization by a genetic algorithm," *IEEE Trans. Antennas Propag.*, vol. 49, pp. 22-31, 2001.

[26] S. Caorsi, A. Massa, and M. Pastorino, "A crack identification microwave procedure based on a genetic algorithm for nondestructive testing," *IEEE Trans. Antennas Propag.*, vol. 49, pp. 1812-1820, 2001.

[27] S. Caorsi, A. Massa, M. Pastorino, M. Raffetto, and A. Randazzo, "Detection of buried inhomogeneous elliptic cylinders by a memetic algorithm," *IEEE Trans. Antennas Propag.*, vol. 51, pp. 2878-2884, 2003.

[28] S. Caorsi, A. Massa, M. Pastorino, and A. Randazzo, "Electromagnetic detection of dielectric scatterers using phaseless synthetic and real data and the memetic algorithm," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, pp. 2745-2753, 2003.

[29] S. Caorsi, A. Massa, M. Pastorino, and A. Rosani, "Microwave medical imaging: Potentialities and limitations of a stochastic optimization technique," *IEEE Trans. Microwave Theory Tech.*, vol. 52, pp. 1909-1916, 2004.

[30] S. Caorsi, A. Massa, M. Pastorino, and M. Donelli, "Improved microwave imaging procedure for nondestructive evaluations of two-dimensional structures," *IEEE Trans. Antennas Propag.*, vol. 52, pp. 1386-1397, 2004.

[31] S. Caorsi, M. Donelli, A. Lommi, and A. Massa, "Location and imaging of two-dimensional scatterers by using a particle swarm algorithm," *J. Electromagn. Waves Appl.*, vol. 18, pp. 481-494, 2004.

[32] S. Caorsi, A. Massa, M. Pastorino, M. Raffetto, and A. Randazzo, "Microwave imaging of cylindrical inhomogeneities based on an analytical forward solver and multiple illuminations," in *Proc. 2004 IEEE Int. Workshop on Imaging Systems and Techniques*, Stresa, Italy, 2004, pp. 100-105.

[33] U. K. Chakraborty, *Advances in Differential Evolution*. Berlin Heidelberg: Springer, 2008.

[34] X. Chen, K. O'Neill, B. E. Barrowes, T. M. Grzegorczyk, and J. A. Kong, "Application of a spheroidal-mode approach and a differential evolution algorithm for inversion of magneto-quasistatic data in UXO discrimination," *Inverse Problems*, vol. 20, pp. 27-40, 2004.

[35] W. C. Chew, *Waves and Fields in Inhomogeneous Media*. New York, NY: IEEE Press, 1995.

[36] W. C. Chew, J.-M. Jin, E. Michielssen, and J. Song, *Fast and Efficient Algorithms in Computational Electromagnetics*. Artech House, Boston, 2001.

[37] W. Chien and C.-C. Chiu, "Using NU-SSGA to reduce the searching time in inverse problem of a buried metallic object," *IEEE Trans. Antennas Propag.*, vol. 53, pp. 3128-3134, 2003.

[38] W. Chien and C.-C. Chiu, "Using NU-SSGA to reduce the searching time in inverse problem of a buried metallic object," *IEEE Trans. Antennas Propag.*, vol. 53, pp. 3128-3134, 2005.

[39] C.-C. Chiu and P.-T. Liu, "Image reconstruction of a perfectly conducting cylinder by the genetic algorithm," *IEE Proc. Microw. Antennas Propag.*, vol. 143, pp. 249-253, 1996.

[40] C. C. Chiu and P. T. Liu, "Electromagnetic transverse electric-wave inverse scattering of a conductor by the genetic algorithm," *Int. J. Imaging Syst. Technol.*, vol. 9, pp. 388-394, 1998.

[41] C.-C. Chiu and W.-T. Chen, "Electromagnetic imaging for an imperfectly conducting cylinder by the genetic algorithm," *IEEE Trans. Microwave Theory Tech.*, vol. 48, pp. 1901-1905, 2000.

[42] M. Clerc and J. Kennedy, "The particle swarm - Explosion, stability, and convergence in a multidimensional complex space," *IEEE Trans. Evol. Comput.*, vol. 6, pp. 58-73, 2002.

[43] C. Coello, D. A. Van Veldhuizen, G. B. Lamont, *Evolutionary Algorithms for Solving Multi-Objective Problems* (2nd Ed.). New York: Springer, 2007.

[44] A. Collaro, G. Franceschetti, F. Palmieri, and M. S. Ferreiro, "Phase unwrapping by means of genetic algorithms," *Journ. Opt. Soc. America A*, vol. 15, pp. 407-418, 1998.

[45] D. Costa and A. Hertz, "Ants can colour graphs," *J. Operat. Res. Soc.*, vol. 48, pp. 295-305, 1997.

[46] R. Dawkins, *The Selfish Gene*. New York: Oxford Univ. Press, 1976.

[47] K. A. De Jong, "An analysis of the behavior of a class of genetic adaptive systems," Ph. D. Dissertation, Univ. Michigan, 1975.

[48] K. A. De Jong and W. M. Spears, "A formal analysis of the role of multi-point crossover in genetic algorithms," *Annals of Mathematics and Artificial Intelligence*, vol. 5, pp. 1-26, 1992.

[49] M. Donelli and A. Massa, "Computational approach based on a particle swarm optimizer for microwave imaging of two-dimensional dielectric scatterers," *IEEE Trans. Microw. Theory Tech.*, vol. 53, pp. 1761-1776, 2005.

[50] M. Donelli, G. Franceschini, A. Martini, and A. Massa, "An integrated multiscaling strategy based on a particle swarm algorithm for inverse scattering problems," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, pp. 298-312, 2006.

[51] M. Donelli, D. Franceschini, P. Rocca, and A. Massa, "Three-dimensional microwave imaging problems solved through an efficient multiscaling particle swarm optimization," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, pp. 1467-1481, 2009.

[52] M. Dorigo, V. Maniezzo, and A. Colorni, "Ant system: optimization by a colony of cooperating agents," *IEEE Trans. Syst. Man and Cybern. B* , vol. 26, pp. 29-41, 1996.

[53] M. Dorigo and L. M. Gambarella, "Ant colony system: A cooperative learning approach to the

traveling salesman problem," *IEEE Trans. Evol. Comput.*, vol. 1, pp. 53-66, 1997.

[54] M. Dorigo, T. Stutzle, *Ant Colony Optimization*. Cambridge, MA: MIT Press, 2004.

[55] M. Dorigo and C. Blum, "Ant colony optimization theory: A survey," *Theoretical Computer Science*, vol. 344, pp. 243-278, 2005.

[56] R. C. Eberhart and Y. Shi, "Comparing inertia weights and constriction factors in particle swarm optimization," in *Proc. Congr. Evolutionary Computation*, Piscataway, NJ, 2000, pp. 84-88.

[57] A. E. Eiben, E. H. L. Aarts, and K. M. Van Hee, "Global convergence of genetic algorithms: A Markov chain analysis," in *Parallel Problem Solving from Nature*, H.-P. Schwefel and R. Manner, Eds. Berlin Heidelberg: Springer, 1991, pp. 4-12.

[58] A. I. El-Gallad, M. El-Hawary, A. A. Sallam, and A. I Kalas, "Swarm intelligence for hybrid cost dispatch problem," in *Proc. Can. Electrical and Computer Engineering Conf.*, vol. 2, 2001, pp. 753-757.

[59] Y. Ermoliev and R. J. B. Wets, *Numerical Techniques for Stochastic Optimization*. New York: Springer, 1988.

[60] H.-Y. Fan and J. Lampinen, "A trigonometric mutation operation to differential evolution," *J. Global Optimization*, vol. 27, pp. 105-129, 2003.

[61] S. Forrest, "Genetic algorithms: Principles of natural selection applied to computation," *Sci.*, vol. 261, pp. 872-878, 1993.

[62] G. Franceschini, M. Donelli, R. Azaro, and A. Massa, "Inversion of phaseless total field data using a two-step strategy based on the iterative multiscaling approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, pp. 3527-3539, 2006.

[63] R. Gamperle, S. D. Muller, and P. Koumoutsakos, "A parameter study for differential evolution," in *Advances in Intelligent Systems, Fuzzy Systems, Evolutionary Computation*, A. Grmela and N. Mastorakis, Eds. WSEAS Press, 2002, pp. 293-298.

[64] S. Garnier, J. Gautrais, and G. Theraulaz, "The biological principles of swarm intelligence," *Swarm Intelligence*, vol. 1, pp. 3-31, 2007.

[65] D. B. Ge, Y. B. Yan, S. Y. Shi, and X. C. Nie, "Genetic algorithm applied to inverse scattering: Two examples," in *Proc. Asia Pacific Microwave Conf.*, pp. 689-692, 1997.

[66] S. Genovesi, E. Salerno, A. Monorchio, and G. Manara, "A permittivity range profile reconstruction technique for multilayered structures by a line process-augmented genetic algorithm," in *Proc. 2005 IEEE AP-S Int. Symp.*, Washington DC, USA, 20-25 Jun. 2005, pp. 216-219.

[67] D. E. Goldberg and P. Segrest, "Finite Markov chain analysis of genetic algorithms," in *Proc. 2nd Int. Conf. Genetic Algorithms*, pp. 1-8, 1987.

[68] D. E. Goldberg, "Simple genetic algorithms and the minimal deceptive problem," in: L. Davis (Ed.), Genetic Algorithms and Simulated Annealing, Pitman, London, UK, 1987, pp. 74-88.

[69] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*. Boston, MA: Addison-Wesley, 1989.

[70] S. Goonatilake and P. Treleaven, *Intelligent Systems for Finance and Business*. Chichester: Wiley, 1995.

[71] P. P. Grassé, "La reconstruction du nid et les coordinations inter-individuelles chez Bellicositermes

Natalensis et Cubitermes sp. La théorie de la stigmergie : essai dâinterprétation du comportement des termites constructeurs," *Insectes Sociaux*, vol. 6, pp. 41-81, 1959.

[72] W. J. Gutjahr, "A graph-based ant system and its convergence," *Future Generat. Comput. Syst.*, vol. 16, pp. 873-888, 2000.

[73] W. J. Gutjahr, "ACO algorithms with guaranteed convergence to the optimal solution," *Informat. Process. Lett.*, vol. 82, pp. 145-153, 2002.

[74] P. G. Harrald, "Evolutionary algorithms and economic models: A view," in *Proc. 5th Annu. Conf. on Evolutionary Programming*, Cambridge, MA: MIT Press, 1996, pp. 3-7.

[75] H. Holland, *Adaptation in Natural and Artificial Systems*. Ann Arbor, MI: Univ. Michigan Press, 1975.

[76] R. L. Haupt, "An introduction to genetic algorithms for electromagnetics," *IEEE Antennas Propag.. Mag.*, vol. 37, pp. 7-15, 1995.

[77] J. H. Holland, "Genetic algorithms," *Sci. Amer.*, vol. 267, pp. 66-72, 1992.

[78] J. Horn, "Finite Markov chain analysis of genetic algorithms with niching," in *Proc. 5th Int. Conf. Genetic Algorithms*, pp. 110-117, 1993.

[79] T. Huang and A. S. Mohan, "Microwave imaging of perfect electrically conducting cylinder by micro-genetic algorithm," in *Proc. 2004 IEEE AP-S Int. Symp.*,Monterey, CA, USA , 2004, pp. 221-224.

[80] T. Huang and A. S. Mohan, "Application of particle swarm optimization for microwave imaging of lossy dielectric objects," in *Proc. 2005 IEEE AP-S Int. Symp.*, Washington DC, USA, 20-25 June, 2005, pp. 852-855.

[81] T. Huang and A. S. Mohan, "A hybrid boundary condition for robust particle swarm optimization," *IEEE Antennas Wireless Propagat. Lett.*, vol. 4, pp. 112-117, 2005.

[82] T. Huang and A. S. Mohan, "Significance of neighborhood topologies for the reconstruction of microwave images using particle swarm optimization," in *Proc. APMC-2005 Asia-Pacific Microwave Conference*, 4-7 Dec. 2005, Suzhou, China, vol. 1.

[83] T. Huang and A. S. Mohan, "A microparticle swarm optimizer for the reconstruction of microwave images," *IEEE Trans. Antennas Propag.*, vol. 55, pp. 568-576, 2007.

[84] C. Kehyayan, N. Mansour, and H. Khachfe, "Evolutionary algorithm for protein structure prediction," in *Proc. ICACTE '08 Int. Conf. Advanced Computer Theory and Engineering*, 25-27 Sept. 2008, Cairo, Egypt, pp. 925-929.

[85] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. 1995 IEEE Intern. Conf. Neural Networks*, Piscataway, NJ (USA), 1995, pp. 1942-1948.

[86] J. Kennedy and R. C. Eberhart, "A discrete binary version of the particle swarm algorithm," in *Proc. World Multiconf. Syst., Cyber. Informatics*, Piscataway, NJ (USA), 1997, pp. 4104-4109.

[87] J. Kennedy, R. C. Eberhart, and Y. Shi, *Swarm Intelligence*. San Francisco, CA: Morgan Kaufmann, 2001.

[88] J. R. Koza, *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. Cambridge, MA: The MIT Press, 1992.

[89] C. Z. Janikow and Z. Michalewicz, "An experimental comparison of binary and floating point representations in genetic algorithms," in *Proc. 4th Int. Conf. Genetic Algorithms*, San Diego,

CA, USA, Jul. 1991, pp. 31-36.

[90] J. M. Johnson and Y. Ramat-Samii, "Genetic algorithms in engenering electromagnetics," *IEEE Antennas Propag. Mag.*, vol. 39, pp. 7-21, 1997.

[91] R. L. Johnston, *Applications of Evolutionary Computation in Chemistry.* Berlin Heidelberg: Springer, 2004.

[92] A. Lewandowski and V. L. Volkovich, *Multiobjective Problems of Mathematical Programming.* Eds. Spinger 1991, *Proc. Int. Conf. Multiobjective Problems of Mathematical Programming*, Yalta, USSR, Oct. 26-Nov. 2, 1988.

[93] F. Li, X. Chen, and K. Huang, "Microwave imaging a buried object by the GA and using the s11 parameter," *Progress In Electromagnetics Research*, vol. 85, pp. 289-302, 2008.

[94] J. Liu and J. Lampinen, "On setting the control parameter of the differential evolution method," in *Proc. 8th Int. Conf. Soft Computing (MENDEL 2002)*, Brno, Czech Republic, 2002, pp. 11-18.

[95] V. Maniezzo, "Exact and approximate nondeterministic tree-search procedures for the quadratic assignment problem," *INFORMS J. Comput.*, vol. 11, pp. 358-369, 1999.

[96] R. Marklein, K. Balasubramanian, A. Qing, and K. J. Langenberg, "Linear and nonlinear iterative scalar inversion of multi-frequency multi-bistatic experimental electromagnetic scattering data," *Inverse Problems*, vol. 17, pp. 1597-1610, 2001.

[97] A. Massa, M. Pastorino, and A. Randazzo, "Reconstruction of two-dimensional buried objects by a differential evolution method," *Inverse Problems*, vol. 20, pp. 135-150, 2004.

[98] A. Massa, D. Franceschini, G. Franceschini, M. Pastorino, M. Raffetto, and M. Donelli, "Parallel GA-based approach for microwave imaging applications," *IEEE Trans. Antennas Propag.*, vol. 53, pp. 3118-3127, 2005.

[99] R. Mendes, J. Kennedy, and J. Neves, "Watch thy neighbor or how the swarm can learn from its environment," in *Proc. IEEE Swarm Intell. Symp.*, pp. 88-94, 2003.

[100] Z. Q. Meng, T. Takenaka, and T. Tanaka, "Image reconstruction of two-dimensional impenetrable objects using genetic algorithm," *J. Electromag. Waves Appl.*, vol. 13, pp. 95-118, 1999.

[101] Z.-Q. Meng, T. Takenaka, and S. He, "A genetic algorithm with an adaptive chromosome structure for reconstruction of radome parameters using a Gaussian beam," *Microw. Opt. Technol. Lett.*, vol. 25, pp. 323-327, 2000.

[102] D. Merkle and M. Middendorf, "Ant colony optimization with global pheromone evaluation for scheduling a single machine," *Appl. Intell.*, vol. 18, pp. 105-111, 2003.

[103] P. Merz and B. Freisleben, "Fitness landscape analysis and memetic algorithms for the quadratic assignment problem," *IEEE Trans. Evol. Comput.*, vol. 4, pp. 337-352, 2000.

[104] N. Meuleau and M. Dorigo, "Ant colony optimization and stochastic gradient descent," *Artif. Life*, vol. 8, pp. 103-121, 2002.

[105] K. A. Michalski, "Electromagnetic imaging of circular-cylindrical conductors and tunnels using a differential evolution algorithm," *Microw. Opt. Technol. Lett.*, vol. 27, pp. 330-334, 2000.

[106] K. A. Michalski, "Electromagnetic imaging of elliptical-cylindrical conductors and tunnels using a differential evolution algorithm," *Microw. Opt. Technol. Lett.*, vol. 28, pp. 164-169, 2001.

[107] S. M. Mikki and A. A. Kishk, "Physical theory for particle swarm optimization," *Progr.*

Electromagn. Wave Res., vol. 75, pp. 171-207, 2007.

[108] A. F. Morabito, I. Catapano, M. D'Urso, L. Crocco, and T. Isernia, "A stepwise approach for quantitative 3D imaging: Rationale and experimental results," in *Proc. 2009 IEEE AP-S Int. Symp.*, Charleston, SC, 2009, 112.1.

[109] P. Moscato, "On evolution, search, optimization, GAs and martial arts: Optimization towards memetic algorithms," California Inst. Technol., Pasadena, CA, Tech. Rep. Caltech Concurrent Comput. Prog. Rep. 826, 1989.

[110] T. Niwa and M. Tanaka, "On the mean convergence time for simple genetic algorithms," in *Proc. 1995 IEEE Int. Conf. Evolutionary Computation*, vol. 1, pp. 373-377.

[111] J. Nocedal and S. J. Wright, *Numerical Optimization*. New York: Springer, 1999.

[112] E. Ozcan and C. K. Mohan, "Particle swarm optimization: Surfing the waves," in *Proc. 1999 Congr. Evolutionary Computation*, Washington, DC, July 1999, pp. 1939-1944.

[113] V. Pareto, *Cours d'Economie Politique*. Lousanne: F. Rouge, 1986.

[114] C. S. Park and B. S. Jeong, "Reconstruction of a high contrast and large object by using the hybrid algorithm combining a Levenberg-Marquardt algorithm and a genetic algorithm," *IEEE Trans. Mag.*, vol. 35, pp. 1582-1585, 1999.

[115] M. Pastorino, A. Massa, and S. Caorsi, "A microwave inverse scattering technique for image reconstruction based on a genetic algorithm," *IEEE Trans. Instrum. Meas.*, vol. 49, pp. 573-578, 2000.

[116] M. Pastorino, S. Caorsi, and A. Massa, "A global optimization technique for microwave nondestructive evaluation," *IEEE Trans. Instrum. Meas.*, vol. 51, pp. 666-673, 2002.

[117] M. Pastorino, S. Caorsi, A. Massa, and A. Randazzo, "Reconstruction algorithms for electromagnetic imaging," *IEEE Trans. Instrum. Meas.*, vol. 53, pp. 692-699, 2004.

[118] M. Pastorino, "Stochastic optimization methods applied to microwave imaging: A review," *IEEE Trans. Antennas Propag.*, vol. 55, pp. 538-548, 2007.

[119] E. Polak, *Computational Methods in Optimization*. New York: Academic Press, 1971.

[120] R. Poli, J. Kennedy, and T. Blackwell, "Particle swarm optimization - An overview," *Swarm Intelligence*, vol. 1, pp. 33-57, 2007.

[121] K. V. Price, R. M. Storn, and J. A. Lampinen, *Differential Evolution: A Practical Approach to Global Optimization*. Berlin Heidelberg: Springer, 2005.

[122] X. Qi and F. Palmieri, "Theoretical analysis of evolutionary algorithms with an infinite population size in continuous space - Part I: Basic properties of selection and mutation," *IEEE Trans. Neural Networks*, vol. 5, pp. 102-119, 1994.

[123] X. Qi and F. Palmieri, "Theoretical analysis of evolutionary algorithms with an infinite population size in continuous space - Part II: Analysis of the diversification role of crossover," *IEEE Trans. Neural Networks*, vol. 5, pp. 120-129, 1994.

[124] A. Qing, C. K. Lee, and L. Jen, "Microwave imaging of parallel perfectly conducting cylinders using real-coded genetic algorithm," *J. Electromag. Waves Appl.*, vol. 13, pp. 1121-1143, 1999.

[125] A. Qing and C. K. Lee, "Microwave imaging of a perfectly conducting cylinder using a real-coded genetic algorithm," in *Proc. IEE Microwave Antennas Propag.*, vol. 146, pp. 421-425, 1999.

[126] A. Qing and C. K. Lee, "Microwave imaging of parallel perfectly conducting cylinders using real-

coded genetic algorithm coupled with Newton-Kantorovitch method," *J. Electromagn. Waves Appl.*, vol. 14, pp. 799-800, 2000.

[127] A. Qing, C. K. Lee, and L. Jen, "Electromagnetic inverse scattering of two-dimensional perfectly conducting objects by real-coded genetic algorithm," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, pp. 665-676, 2001.

[128] A. Qing, "An experimental study on electromagnetic inverse scattering of a perfectly conducting cylinder by using the real-coded genetic algorithm," *Microw. Opt. Technol. Lett.*, vol. 30, pp. 315-320, 2001.

[129] A. Qing, "Electromagnetic imaging of two-dimensional perfectly conducting cylinders with transverse electric scattered field," *IEEE Trans. Antennas Propag.*, vol. 50, pp. 1786-1794, 2002.

[130] A. Qing, "Electromagnetic inverse scattering of multiple two-dimensional perfectly conducting objects by the differential evolution strategy," *IEEE Trans. Antennas Propag.*, vol. 51, pp. 1251-1262, 2003.

[131] A. Qing, "Electromagnetic inverse scattering of multiple perfectly conducting cylinders by differential evolution strategy with individuals in groups (GDES)," *IEEE Trans. Antennas. Propag.*, vol. 52, pp. 1223-1229, 2004.

[132] A. Qing, "Dynamic differential evolution strategy and applications in electromagnetic inverse scattering problems," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, pp. 116-125, 2006.

[133] A. Qing, "A parametric study on differential evolution based on benchmark electromagnetic inverse scattering problem," in *Proc. 2007 IEEE Congress Evol. Comput. (CEC 2007)*, Tokyo, Japan, pp. 1904-1909.

[134] I. T. Rekanos, "Shape reconstruction of a perfectly conducting scatterer using differential evolution and particle swarm optimization," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, pp. 1967-1974, 2008.

[135] M. Reimann, K. Doerner, and R. F. Hartl, "D-ants: Savings based ants divide and conquer the vehicle routing problem," *Computers & Operations Research*, vol. 31, no. 4, pp. 563-591, 2004.

[136] M. Reimann, *Progress in Complex Systems Optimization*. Berlin Heidelberg: Springer, 2007.

[137] J. H. Richmond, "Scattering by a dielectric cylinder of arbitrary cross-section shape," *IEEE Trans. Antennas Propag.*, vol. 13, pp. 334-341, 1965.

[138] J. Robinson and Y. Rahmat-Samii, "Particle swarm optimization in electromagnetics," *IEEE Trans. Antennas Propag.*, vol. 52, pp. 397-407, 2004.

[139] G. Rudolph, "Convergence analysis of canonical genetic algorithms," *IEEE Trans. Neural Networks*, vol. 5, pp. 96-101, 1994.

[140] A. Sabouni, M. Xu, S. Noghanian, P. Thulasiraman, and S. Pistorius, "Efficient microwave breast imaging technique using parallel finite difference time domain and parallel genetic algorithms," in *Proc. 2007 IEEE AP-S Int. Symp.*, Albuquerque, NM, 2007, pp. 2176-2179.

[141] A. Saeedfar and K. Barkeshli, "Shape reconstruction of three-dimensional conducting curved plates using physical optics, NURBS modeling, and genetic algorithm," *IEEE Trans. Antennas Propag.*, vol. 54, pp. 2497-2507, 2006.

[142] J. D. Schaffer, "Multiple objective optimization with vector evaluated genetic algorithms," Ph.D.

dissertation, Vanderbilt Univ., Nashville, TN, 1984.

[143] J . D. Schaffer, "Multiple objective optimization with vector evaluated genetic al gorithms," in *Proc. 1st Int. Conf. Genetic Algorithms*, Hillsdale, NJ, Lawrence Erlbaum, 1985, pp. 93-100.

[144] J. J. Schneider and S. Kirkpatrick, *Stochastic Optimization*. Berlin Heidelberg: Springer, 2006.

[145] A. Semnani and M. Kamyab, "An enhanced method for inverse scattering problems using Fourier series expansion in conjunction with FDTD and PSO," *Progress In Electromagnetics Research*, vol. 76, pp. 45-64, 2007.

[146] Y. Shi and R. C. Eberhart, "A modified particle swarm optimizer," in *Proc. 1998 IEEE Inter. Conf. Evol. Comput.*, IEEE Press, Piscataway, NJ, pp. 69-73.

[147] Y. Shi and R. C. Eberhart, "Empirical study of particle swarm optimization," in *Proc. Congr. Evolutionary Computation*, DC, 1999, pp. 1945-1950.

[148] K. Socha, "ACO for continuous and mixed-variable optimization," in *Proc. 4th Int. Workshop Ant Colony Optim. Swarm Intell.* (ANTS'2004), Brussels, Belgium, Sep. 5-8, 2004.

[149] K. Socha and C. Blum, "An ant colony optimization algorithm for continuous optimization: application to feed-forward neural network training," *Neural Comput. & Applic.*, vol. 16, pp. 235-247, 2007.

[150] R. Storn and K. Price, "Differential evolution - A simple and efficient adaptive scheme for global optimization over continuous spaces," International Comput. Science Inst., Berkeley, CA, Tech. Rep. TR-95-012, 1995

[151] R. Storn and K. Price, "Differential evolution - A simple and efficient heuristic for global optimization over continuous spaces," *J. Global Optim.*, vol. 11, pp. 341-359, 1997.

[152] T. Stutzle and M. Dorigo, "A short convergence proof for a class of ant colony optimization algorithms," *IEEE Trans. Evol. Comput.*, vol. 6, pp. 358-365, 2002.

[153] J. Suzuki, "A Markov chain analysis on a genetic algorithm," in *Proc. 5th Inter. Conf. Genetic Algorithms*, University of Illinois at Urbana Champaign, pp. 146-153, 1993.

[154] T. Takenaka, Z. Q. Meng, T. Tanaka, and W. C. Chew, "Local shape function combined with genetic algorithm applied to inverse scattering for strips," *Microw. Opt. Technol. Lett.*, vol. 16, pp. 337-341, 1997.

[155] G. Theraulaz and E. Bonabeau, "A brief history of stigmergy," *Artificial Life*, vol. 5, pp. 97-116, 1999.

[156] X. L. Travassos, D. A. G. Vieira, N. Ida, C. Vollaire, and A. Nicolas, "Inverse algorithms for the GPR assessment of concrete structures," *IEEE Trans. Magn.*, vol. 44, pp. 994-997, 2008.

[157] I. C. Trelea, "The particle swarm optimization algorithm: convergence analysis and parameter selection," *Information Processing Letters*, vol. 85, pp. 317-325, 2003.

[158] G. N. Vanderplaats, *Numerical optimization techniques for engineering design: With applications*. New York: McGraw-Hill, 1984.

[159] M. D. Vose and G. E. Liepins, "Punctuated equilibria in genetic search," *Complex Systems*, vol. 5, pp. 31-44, 1991.

[160] W. Wang, Y. Lu, J. Fu, and Y. Z. Xiong, "Particle swarm optimization and finite-element based approach for microwave filter design," *IEEE Trans. Magnetics*, vol. 41, pp. 1800-1803, 2005.

[161] D. S. Weile and E. Michielssen, "Genetic algorithm optimization applied to electromagnetics: A

review," *IEEE Trans. Antennas Propag.*, vol. 45, pp. 343-353, 1997.

[162] W. Weishui and X. Chen, "Convergence theorem of genetic algorithm," in *Proc. 1996 IEEE Int. Conf. Systems, Man, and Cybernetics*, Beijing, China, pp. 1676-1681.

[163] R. A. Wildman and D. S. Weile, "Geometry reconstruction of conducting cylinders using genetic programming," *IEEE Trans. Antennas Propag.*, vol. 55, pp. 629-636, 2007.

[164] R. A. Wildman and D. S. Weile, "Greedy search and a hybrid local optimization/genetic algorithm for tree-based inverse scattering," *Microw. Opt. Technol. Lett.*, vol. 50, pp. 822-825, 2008.

[165] D. H. Wolpert and W. G. Macready, "No free lunch theorems for optimization," *IEEE Trans. Evol. Comput.*, vol. 1, pp. 67-82, 1997.

[166] F. Xiao and H. Yabe, "Microwave imaging of perfectly conducting cylinders from real data by micro genetic algorithm couple with deterministic method," *IEICE Trans. Electron.*, vol. E81-C, pp. 1784-1792, 1998.

[167] M. Xu, A. Sabouni, P. Thulasiraman, S. Noghanian, and S. Pistorius, "Image reconstruction using microwave tomography for breast cancer detection on distributed memory machine," in *Proc. 2007 Int. Conf. Parallel Processing (ICPP 2007)*, Xi'an, China, September 10-14, 2007, p. 36.

[168] F. Xue, A. C. Sanderson, and R. J. Graves, "Multi-objective differential evolution - Algorithm, convergence analysis, and applications," in *Proc. 2005 IEEE Congress on Evolutionary Computation*, 2-4 Sep. 2005, Edinburgh, UK, vol. 1, pp. 743-750.

[169] S.-Y. Yang, H.-K. Choi, J.-W. Ra, "Reconstruction of a large and high-contrast penetrable object by using the genetic and Levenberg-Marquardt algorithms," *Microw. Opt. Tech. Lett.*, vol. 16, pp. 17-21, 1997.

[170] S. H. Zainud-Deen, W. M. Hassen, E. M. Ali, K. H. Awadalla and H. A. Sharshar, "Breast cancer detection using a hybrid finite difference frequency domain and particle swarm optimization techniques," *Progress In Electromagnetics Research B*, vol. 3, pp. 35-46, 2008

[171] D. Zaharie, "Critical values for the control parameters of differential evolution algorithm," in *Proc. 8th Int. Conf. Soft Computing (MENDEL 2002)*, Brno, Czech Republic, 2002, pp. 62-67.

[172] B. Zhao and S. Li, "A convergence proof for ant colony algorithm," in *Proc. 6th World Congress Intelligent Control and Automation*, June 21-23, Dalian, China, vol. 1, pp. 3072-3075, 2006.

[173] Y. Zhou and H. Ling, "Electromagnetic inversion of Ipswich objects with the use of the genetic algorithm," *Microw. Opt. Technol. Lett.*, vol. 33, pp. 457-459, 2002.

[174] Y. Zhou, J. Li, and H. Ling, "Shape inversion of metallic cavities using hybrid genetic algorithm combined with tabu list," *Electron. Lett.*, vol. 39, pp. 280-281, 2003.

[175] E. Zitzler, K. Deb, L. Thiele, C. Coello, and D. Corne, "Evolutionary multi-criterion optimization," in *Proc. 1st Int. Conf. (EMO'01)*, Springer, Berlin, 2001, vol. 1993.

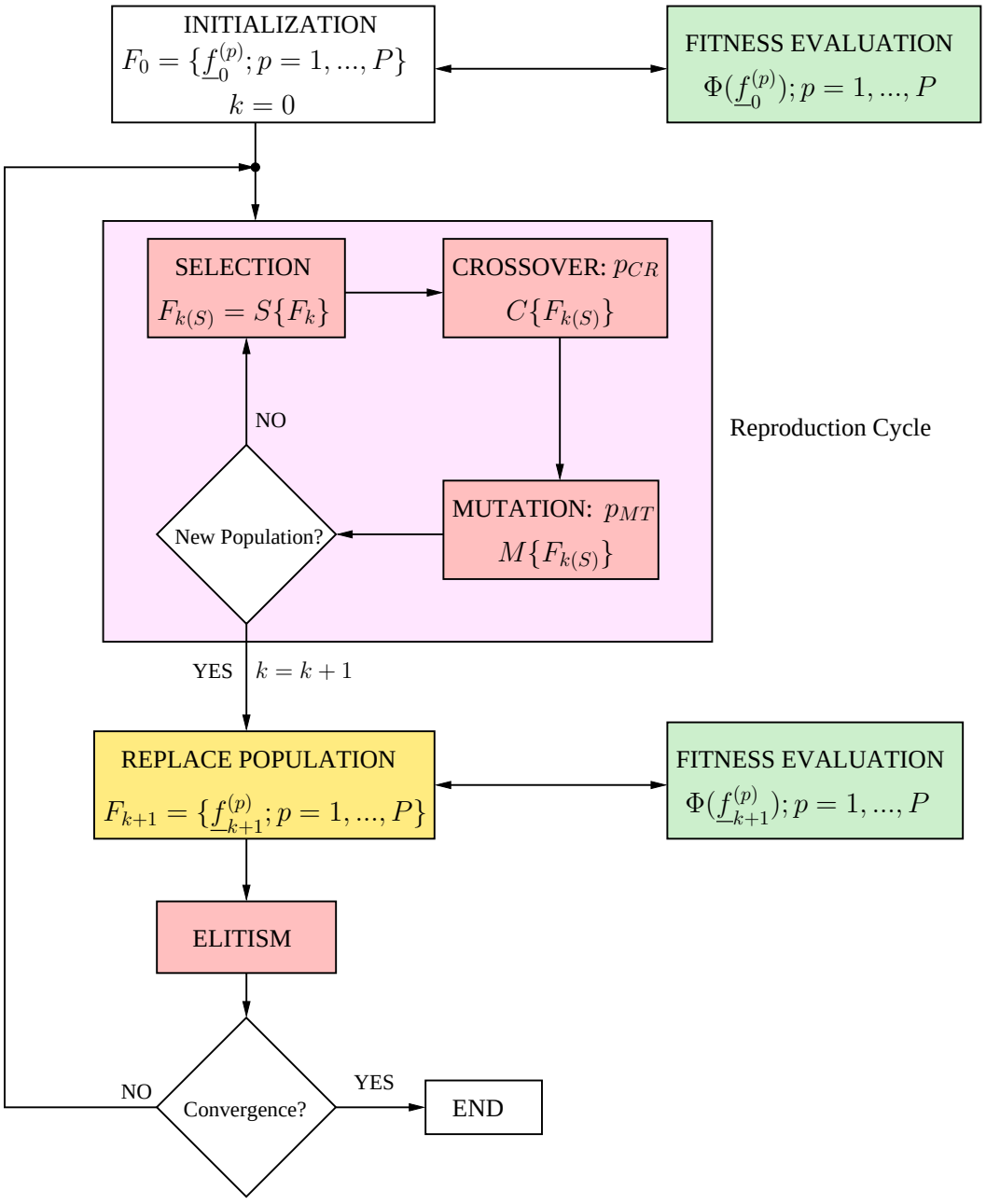

Figure 1. *Genetic Algorithm - Flowchart.*

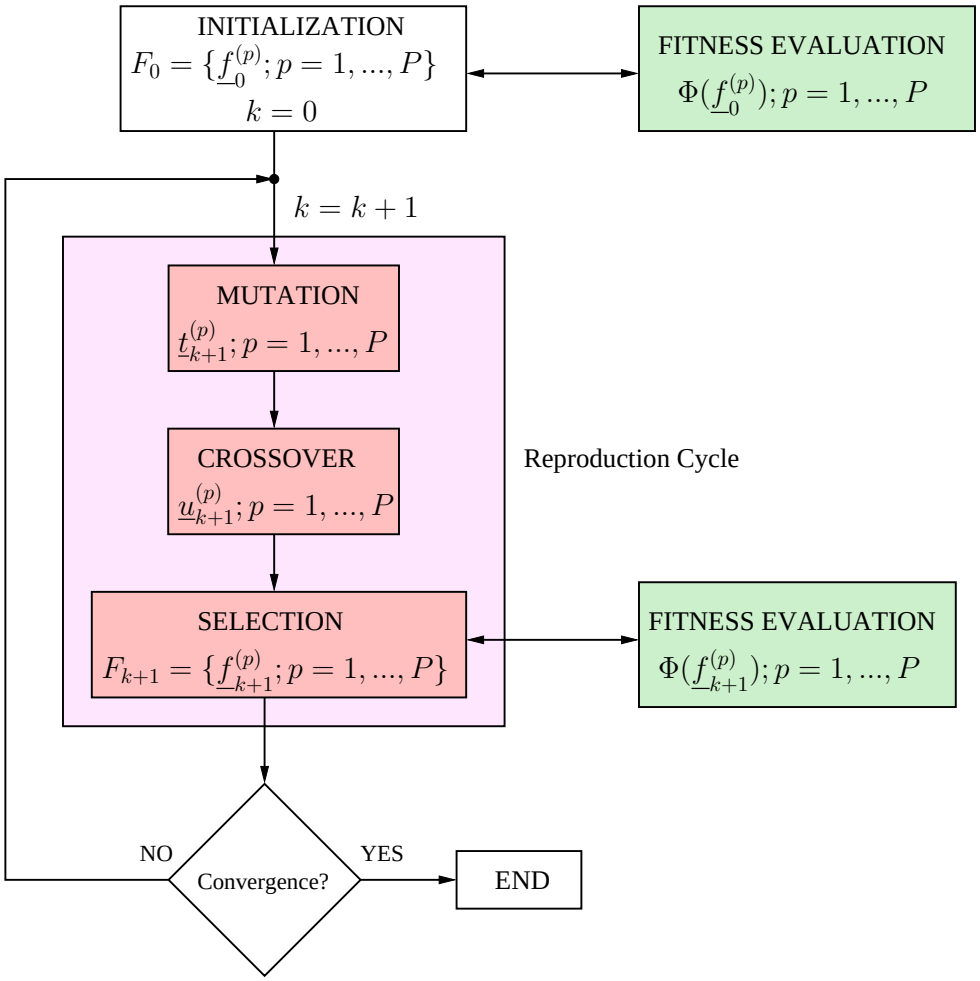

Figure 2. *Differential Evolution - Flowchart.*

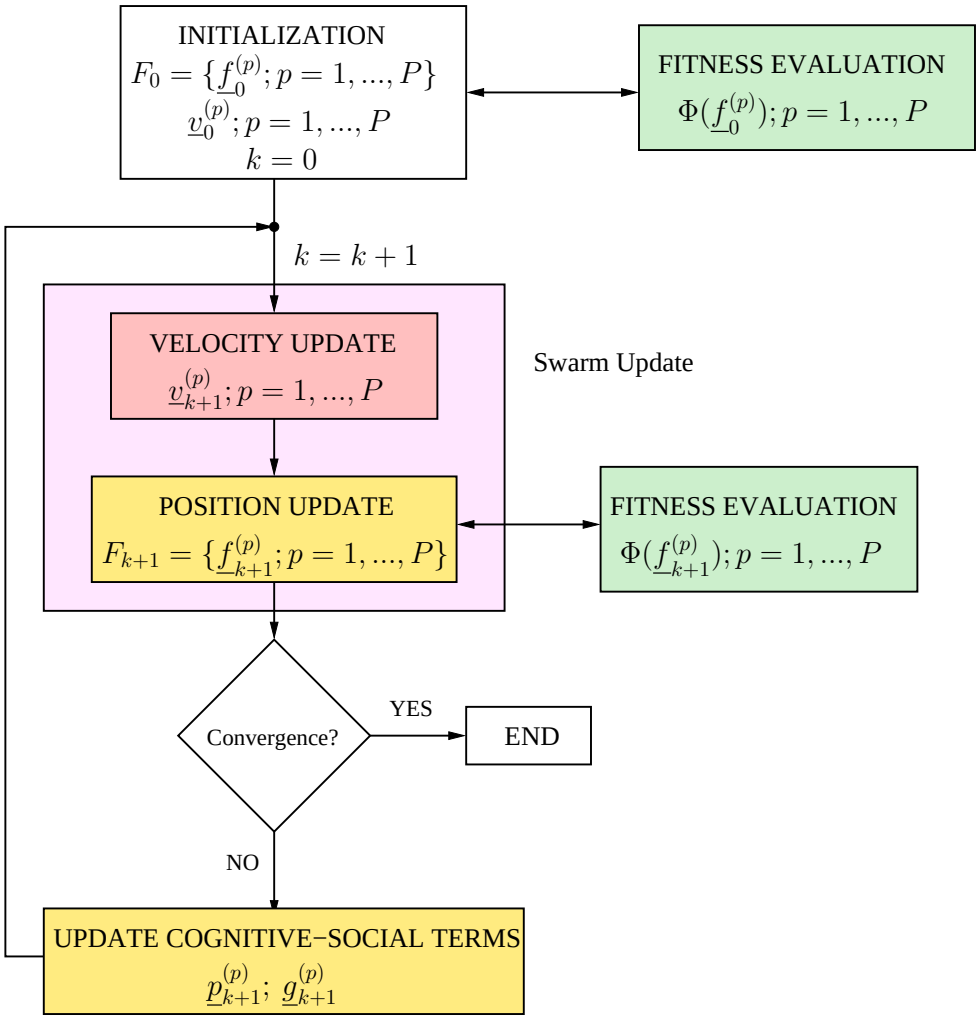

Figure 3. *Particle Swarm Optimizer - Flowchart.*

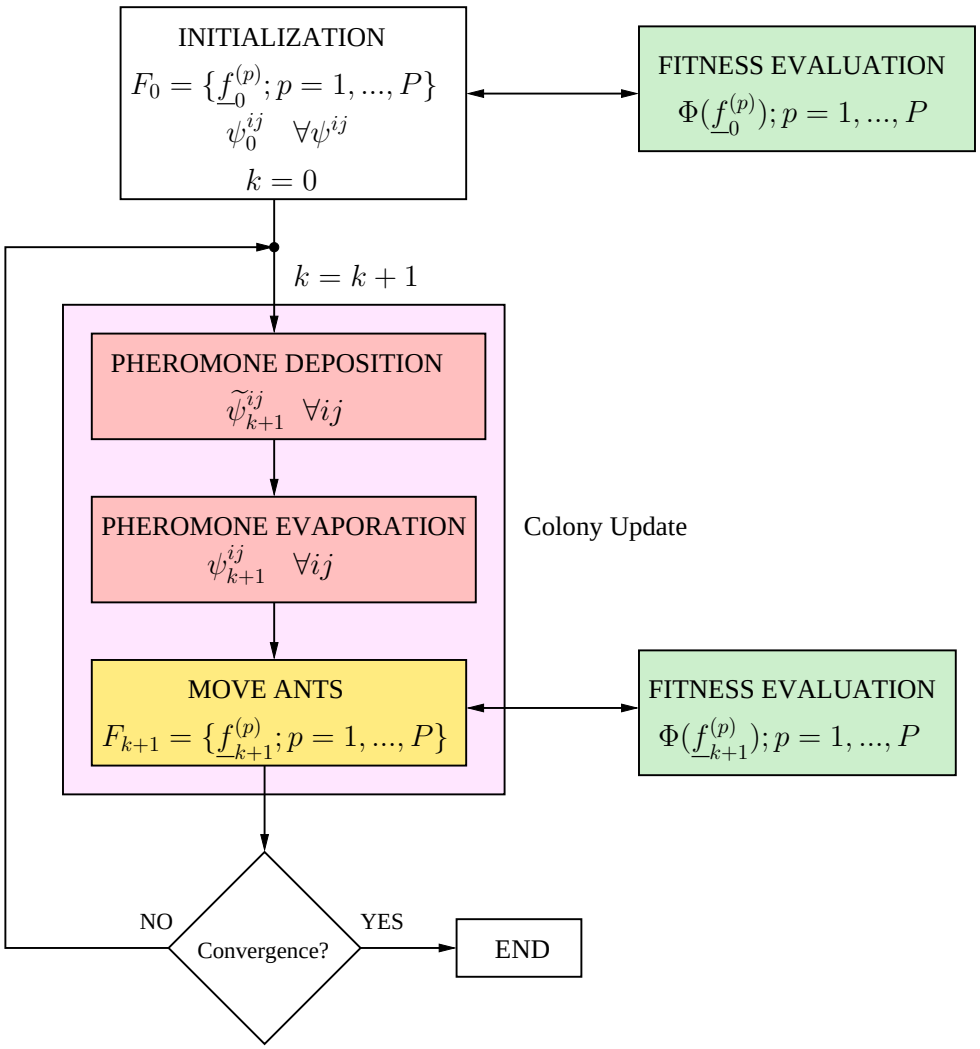

Figure 4. *Ant Colony Optimizer - Flowchart.*

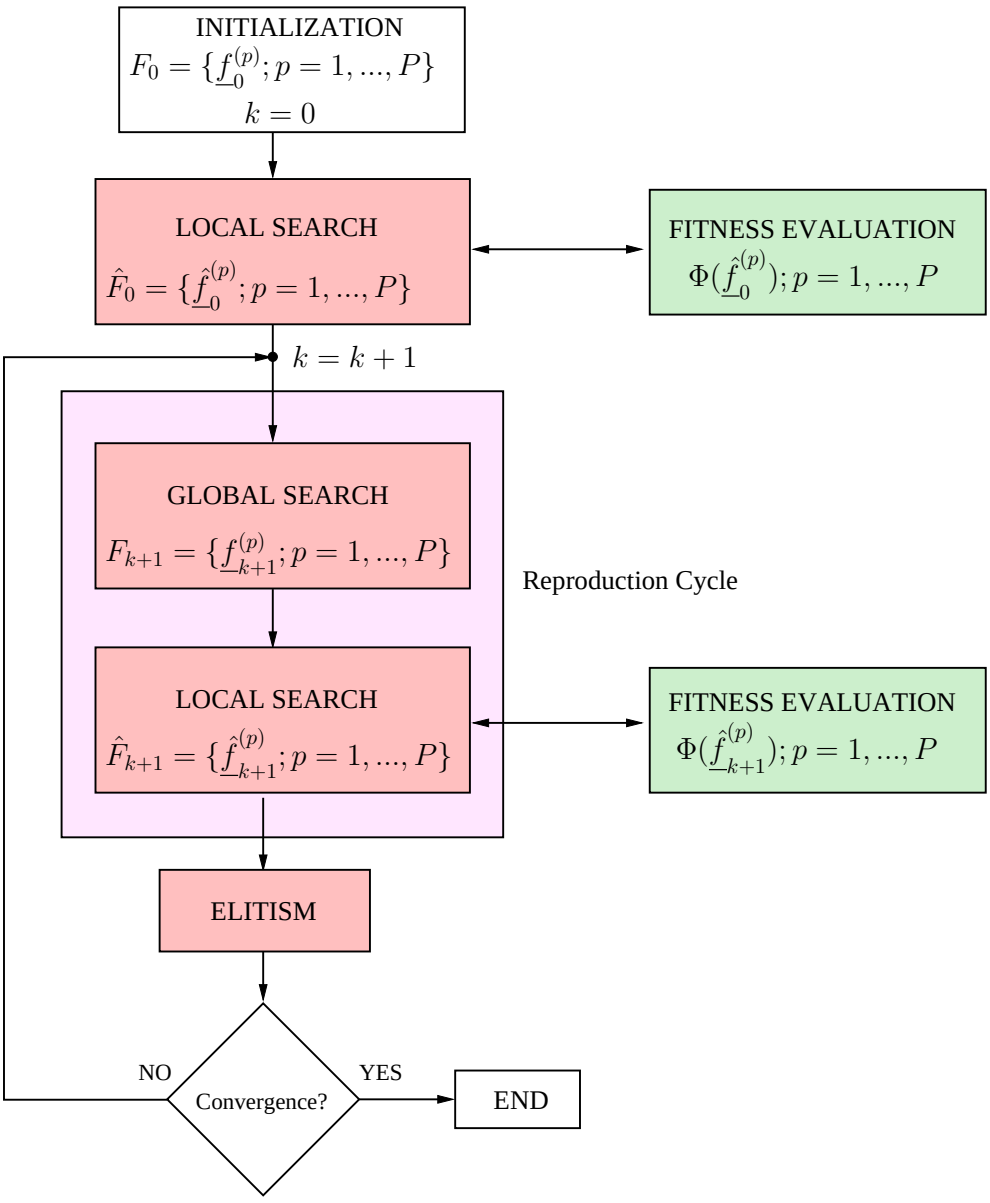

Figure 5. *Memetic Algorithm - Flowchart.*

Algorithm	*Object*	*Geometry*	N	$\frac{P}{N}$	K_{end}	$Time\,[sec]$	PC	$Reference$
BGA	PEC	$2D$	9	~ 33	–	$\sim 1.8 \times 10^3$	$Sun\,Sparc\,20$	[39]
BGA	PEC	$2D$	36	~ 1.4	173	330	$Sun\,Sparc\,20$	[100]
BGA	$PEC*$	$2D$	400	0.25	–	–	–	[173]
BGA	PEC	$3D$	6	–	30	$\sim 6.7 \times 10^4$	$Pentium\,4$	[141]
RGA	PEC	$2D$	8	32	35	$\sim 1.6 \times 10^4$	$IBM\,P-133$	[127]
RGA	$PEC*$	$2D$	6	50	40	$\sim 8.6 \times 10^4$	$MIPS\,R10K$	[96]
μGA	$PEC*$	$2D$	10	–	–	$\sim 1.7 \times 10^4$	$Sun\,Sparc\,20$	[166]
μGA	PEC	$2D$	5	1	1000	$\sim 3.5 \times 10^3$	$Pentium\,4$	[79]
GA/CG	$PEC*$	$2D$	5	40	220	–	–	[174]
$GA/CG + tabu$	$PEC*$	$2D$	5	40	75	–	–	[174]
DE	PEC	$2D$	5	5	40	–	–	[106]
DE	$PEC*$	$2D$	6	~ 7	–	–	$HP\,OmniBook\,XE3$	[130]
DE	PEC	$2D$	16	10	~ 160	$\sim 4 \times 10^3$	$HP\,OmniBook\,XE3$	[130]
$DE\,(GDES)$	PEC	$2D$	16	~ 18	23	$\sim 1.2 \times 10^3$	$Pentium\,4$	[131]
PSO	PEC	$2D$	10	30	200	–	–	[134]
BGA	*conductor*	$2D$	10	30	–	–	–	[41]
RGA	*dielectric*	$2D$	6	50	20	$\sim 8.6 \times 10^4$	$MIPS\,R10K$	[96]
DE	*dielectric*	$2D$	5	5	30	–	–	[106]

Table I. *Computational indexes of EAs applied to the shape reconstruction of metallic and homogeneous dielectric objects. N is the number of unknown parameters, P is the number of trial solutions for each iteration, K_{end} is the number of iteration at convergence. The asterisk $*$ means inversion of experimental data.*

Algorithm	*Object*	*Geometry*	N	$\frac{P}{N}$	K_{end}	*Time* [*sec*]	*PC*	*Reference*
BGA	*dielectric*	$2D$	900	–	2000	270	*Pentium II*	[23]
BGA	*dielectric*	$2D$	400	0.15	8000	–	–	[25]
RGA	*dielectric*	$2D$	810	0.2	10000	$\sim 10^4$	*Pentium*	[24]
RGA	*dielectric*	$2D$	500	–	8000	$\sim 1.2 \times 10^3$	–	[50]
GA/CG	*dielectric*	$2D$	810	0.2	3400	$\sim 2 \times 10^4$	*Pentium*	[24]
PSO	*dielectric*	$2D$	500	0.05	8000	$\sim 7.6 \times 10^2$	–	[50]
PSO†	*dielectric*∗	$2D$	2664	7.5×10^{-3}	–	$1.75 \times K_{end}$	–	[50]
micro − *PSO*	*dielectric*	$2D$	125	0.04	500	–	–	[83]
PSO†	*dielectric*	$3D$	3250	~ 0.04	5100	$\sim 2.1 \times 10^4$	–	[51]

Table II. *Computational indexes of EAs applied to the reconstruction of the dielectric distribution. N is the number of unknown parameters, P is the number of trial solutions for each iteration, K_{end} is the number of iteration at convergence. The asterisk $*$ means inversion of experimental data. The symbol † indicates the use of a multi-resolution strategy.*