Learned Global Optimization for Inverse Scattering Problems - Matching Global Search with Computational Efficiency

M. Salucci, Senior Member, IEEE, L. Poli, Member, IEEE, P. Rocca, Senior Member, IEEE, and A. Massa, Fellow, IEEE

Abstract—The computationally-efficient solution of fully nonlinear microwave inverse scattering problems (ISPs) is addressed. An innovative System-by-Design (SbD) based method is proposed to enable, for the first time to the best of the authors’ knowledge, an effective, robust, and time-efficient exploitation of an evolutionary algorithm (EA) to perform the global minimization of the data-mismatch cost function. According to the SbD paradigm as suitably applied to ISPs, the proposed approach founds on (i) a smart re-formulation of the ISP based on the a-priori information on the imaged targets for defining a minimum-dimensionality and representative set of degrees-of-freedom (DoFs) and on (ii) the artificial-intelligence (AI)-driven integration of a customized global search technique with a digital twin (DT) predictor based on the Gaussian Process (GP) theory. Representative numerical and experimental results are provided to assess the effectiveness and the efficiency of the proposed approach also in comparison with competitive state-of-the-art inversion techniques.

Index Terms—Inverse Scattering (IS), Evolutionary Algorithms (EAs), System-by-Design (SbD), Digital Twin (DT), Artificial Intelligence (AI), Learning-by-Examples (LBE), Gaussian Processes (GPs).

I. INTRODUCTION AND MOTIVATION

In microwave imaging, an electromagnetic (EM) source illuminates an inaccessible investigation domain to be non-invasively reconstructed by inverting the scattered field data collected in an external observation domain [1]. Depending on the application at hand, both qualitative (i.e., detection, localization, and shaping) and quantitative (i.e., EM properties characterization) reconstructions can be yielded by solving an inverse scattering problem (ISP). ISPs arise in free-space imaging, biomedical diagnostics [2]-[6], subsurface and ground penetrating radar (GPR) investigations [7]-[10], non-destructive testing and evaluation (NDT/NDE) [11]-[14], and through-the-wall imaging (TWI) [15]-[17]. Recently, microwave imaging techniques, based on inverse scattering (IS) formulations, have been also successfully applied to innovative contexts such as, for instance, food quality assessment [18]-[20]. However, solving an ISP is not a trivial task and it poses several challenges due to the intrinsic complexity of the scattering phenomena in the microwave regime described by the Maxwell’s equations. First, the non-uniqueness of the solution, caused by the presence of non-radiating currents induced in the investigation domain, that do not contribute to the scattered data. Second, the non-linearity related to the multiple scattering effects [1]. To properly address such issues for yielding robust/reliable data-inversions, many effective strategies have appeared in the state-of-the-art literature. For instance, Born-based [21] and Rytov-based [22] approximations simplify the IS equations as linearly depending on the unknown contrast distribution. However, they have limited applications to weak scatterers. Otherwise, innovative reformulations of the scattering equations as, for instance, the contraction integral equation (CIE) method, have been introduced to deal with the non-linearity by properly redefining the contrast function [15][23]. Differently, contrast source inversion (CSI) techniques proved to be an effective alternative to the linearization of the data equation [24], even though they are subject to the non-uniqueness of the arising inverse source problem so that multiplicative regularizations have been investigated [25][26]. Regardless of the formulation and unless closed-form solutions, ISPs are generally solved with deterministic (DO) or global (GO) optimization techniques. Strategies belonging to the former class include the subspace optimization method (SOM) [27]-[29], the conjugate gradient (CG) [30], and the inexact Newton method (INM) [31]. To deterministically explore the solution space, these methods typically require the analytic/numerical differentiation of the cost function to be minimized. Consequently, they exhibit a high computational efficiency, but they can be trapped into local-minima/false-
solutions, unless properly initialized within the so-called “attraction basin” of the global optimum.

As for GO methods, nature-inspired strategies (i.e., evolution-ary algorithms (EAs) [32]-[34]) such as genetic algorithms (GAs) [14], particle swarm optimization (PSO) [10], and differential evolution (DE) [35] have been successfully applied to solve ISPs. Thanks to the “hill-climbing” features, they perform an effective global exploration of the solution space by evolving a population of trial solutions with stochastic operators [32] to “escape” from local minima, while converging towards the global optimum. Although successful in several ISP applications and more effective than DOs in sampling nonlinear cost functions, EA-GOs are inherently limited by the computational burden. Indeed, the CPU cost of a stochastic GO is directly linked to the number of agents that evolve throughout the optimization process, which is in turn proportional to the number of degrees-of-freedom (DoFs) that define the dimensionality of the solution space. 1. To partially counteract such a limitation, one practical and effective solution is the integration of EA-GOs with multi-resolution (MR) strategies such as the iterative multi-scaling approach (IMSA) [37]. By adaptively refining the spatial resolution of the reconstruction only within the so-called regions-of-interest (RoI), where the unknown scatterer has been detected, the number of unknowns is strongly reduced at each MR step [10][38] by making computationally-feasible an EA-GO-based optimization.

On the other hand, artificial intelligence (AI) based techniques, belonging to the so-called deep learning (DL) framework [39]-[42], have shown an unprecedented computational efficiency in addressing the pixel-wise inversion of scattered data. However, they still present some unsolved challenges such as the need of huge amounts of training datasets to calibrate thousands of hyper-parameters that define the underlying complex neural network (NN) architecture composed by several hidden layers [39]. Within the AI context, the System-by-Design (SbD) has rapidly emerged as an innovative paradigm for the optimization-driven solution of complex EM problems [43]. The problem at hand is first decomposed into a set of sub-tasks implemented into suitably-defined functional blocks jointly designed with the shared goal of an effective, reliable, and computationally-efficient exploitation of GOs. Such a goal is attained by (i) re-formulating the problem at hand as a GA one described by a minimum-dimensionality set of DoFs enabling a “smart” modeling of the class of targets to be imaged (e.g., piece-wise homogeneous) and (ii) integrating EA-based strategies with fast analysis tools or digital twins (DTs), generated with learning-by-examples (LBEs) techniques [44], to speed up the evaluation (i.e., the cost function computation) of each trial solution. Thanks to its effectiveness and efficiency, the SbD has been already successfully applied to many EM design problems including the synthesis of single radiators [45], wide angle impedance matching layers [46], reflectarrays [47], and meta-material devices [48], but not to ISPs. This paper is then aimed at assessing the SbD in reliably solving fully non-linear ISPs with a computational efficiency, comparable to that of DOs, towards the “holy-grail” of a global real-time optimization.

The paper is organized as follows. The ISP is described and mathematically formulated in Sect. II. Section III details the customization of the SbD paradigm to ISPs and its implementation. Numerical and experimental results are shown in Sect. IV to prove the effectiveness and the efficiency of the proposed method in different operative conditions. Eventually, some conclusions and final remarks are drawn (Sect. V).

II. MATHEMATICAL FORMULATION

Without loss of generality, let us consider a two-dimensional (2D) scenario comprising a square investigation domain D located within a homogeneous, lossless (i.e., conductivity \( \sigma = \sigma_0 = 0 \) [S/m]), and non-magnetic (i.e., permeability \( \mu = \mu_0 \)) background medium of permittivity \( \varepsilon_0 \). By assuming a time-harmonic dependence \( \exp(-j2\pi f t) \), \( f \) being the working frequency, and a transverse magnetic (TM) (i.e., \( z \)-oriented) polarization of the EM field, the scattering phenomena excited by a set of \( V \) monochromatic incident fields, \( \{ T^{(v)}(x, y); \ v = 1, ..., V \} \), which illuminate the investigation domain \( D \), in any \( (x, y) \in D \) are modeled by the following State Equation [1]

\[
T^{(v)}(x, y) = \mathcal{T}^{(v)}(x, y) - \int_D G(x, y, x', y') \mathcal{J}^{(v)}(x', y') \, dx' \, dy'
\]

where

\[
\mathcal{J}^{(v)}(x, y) = \tau(x, y) T^{(v)}(x, y)
\]

is the \( v \)-th \( (v = 1, ..., V) \) equivalent current induced within \( D \), \( \mathcal{T}^{(v)}(x, y) \) is the total field, and

\[
\tau(x, y) = [\varepsilon_r(x, y) - 1] + j \frac{\sigma(x, y)}{2\pi f \varepsilon_0}
\]

is the contrast function that mathematically models the presence, within \( D \), of an unknown scatterer with support \( \Omega \) (i.e., \( \tau(x, y) \neq 0 \) when \( (x, y) \in \Omega \)) whose relative permittivity and conductivity distributions are equal to \( \varepsilon_r(x, y) \) \( \varepsilon_r(x, y) \approx \frac{\varepsilon(x, y)}{\varepsilon_0} \) and \( \sigma(x, y) \), respectively. Moreover,

\[
G(x, y, x', y') = \frac{k_0}{4} \mathcal{H}^{(1)}_{00}(k_0 \sqrt{(x-x')^2 + (y-y')^2})
\]

is the 2D Green’s function of the background medium, \( \mathcal{H}^{(1)}_{00}(\cdot) \) being the zero-th order Hankel’s function of the first kind, and \( k_0 \) is the wavenumber \( k_0 = 2\pi f \sqrt{\varepsilon_0 \mu_0} \).

Otherwise, the EM interactions in the external observation domain \( O \notin D \) \( (O \cap D = \{0\}) \) [1] are described by the Data Equation

\[
S^{(v)}(x, y) = \int_D G(x, y, x', y') \mathcal{J}^{(v)}(x', y') \, dx' \, dy',
\]

where \( S^{(v)}(x, y) \) \( \Phi^{(v)}(x, y) \) \( \Phi^{(v)}(x, y) \approx \mathcal{T}^{(v)}(x, y) - \mathcal{T}^{(v)}(x, y) \) is the scattered field radiated in free-space by the \( v \)-th \( (v = 1, ..., V) \) equivalent source, \( \mathcal{J}^{(v)}(x, y) \), and embedding the information on the unknown scatterer distribution in \( D \).

1It should be pointed out that the term DoF in this paper refers to the minimum number of independent parameters sufficient to describe the solution with a given precision [36].
To numerically deal with (5), the method-of-moments (MoM) is applied by partitioning $D$ into $N$ square sub-domains, $D_n$ being the $n$-th $(n = 1, ..., N)$ discretization domain $(D = \sum_{n=1}^{N} D_n)$ centered at $(x_n, y_n)$ and using $M$ Dirac’s test functions to sample the scattered field at $M$ locations in $O$, $S^{(v)} = \{S^{(v)}(x_m, y_m) ; m = 1, ..., M\}$. The discrete form of (5) is then derived

$$S^{(v)} = G \mathbf{J}^{(v)}$$

where $\mathbf{J}^{(v)} = \{J^{(v)}(x_n, y_n) ; n = 1, ..., N\}$ and $G$ is the $(M \times N)$ external Green’s matrix whose $(m, n)$-th $(m = 1, ..., M; n = 1, ..., N)$ entry is given by $G_{mn} = \int_{D_n} \int_{D_m} \mathbf{H}_0^{(1)}(\rho_{mn}) d\rho' d'y'$. Thus, the inverse problem at hand can be stated as follows

ISP - Starting from the knowledge of the incident, $\{\mathbf{T}^{(v)}(x_n, y_n) ; n = 1, ..., N\}$, and the scattered, $\{S^{(v)}(x_m, y_m) ; m = 1, ..., M\}$, data samples, determine the contrast function distribution, \{\xi(x_n, y_n) ; n = 1, ..., N\}, by solving (6).

III. ShD-BASED INVERSION METHOD

According to the ShD paradigm, the solution of the ISP relies on the exploitation of four interconnected functional blocks, each performing a specific sub-task (Fig. 1). The design and implementation of each block is strongly correlated to the other ones and it is driven by the following shared goals [43]: (i) to yield an effective and reliable solution of the fully non-linear ISP. From an optimization viewpoint, it means to guarantee the convergence towards the global optimum; (ii) to reduce the computational burden required by a standard non-deterministic exploration of the solution space. In other words, the proposed ShD approach is aimed at overcoming the limitation of DOs, which cannot avoid being trapped into local minima unless properly initialized in the “attraction basin” of the actual-solution/global-optimum, while yielding competitive computational performance in solving the ISP so that the following condition on the required CPU-time holds true

$$\Delta t_{ShD} \simeq \Delta t_{DO} \ll \Delta t_{GO}.$$ 

More specifically, the ShD as applied to ISPs is implemented by defining the following blocks (Fig. 1):

1) Problem Formulation (PF) - This block reformulates the ISP to enable an effective, reliable, and computationally-efficient exploitation of GOs by coding the ISP unknowns into a minimum-dimension (yet highly-flexible) set of $K$ degrees-of-freedom (DoFs), $\xi = \{\xi_k ; k = 1, ..., K\}$, to give a “smart” representation of the solution space. Moreover, it defines a suitable cost function, $\Phi(\xi)$, which quantifies the quality of the solution in terms of data mismatch and it represents the unique link between the computational world and the physical one;

2) Data Computation (DC) - In this block, the set of ShD-DoFs, $\xi$, is mapped into a pixel-based representation of the equivalent currents induced within $D$, $\{\mathbf{j}^{(v)}(v) ; v = 1, ..., V\}$, by means of (2) and (1) to compute, through (6), the scattered field distribution in $O$;

3) Cost Function Evaluation (CFE) - This block efficiently evaluates the cost function with a computationally-fast digital twin (DT) [44] of the accurate, but time-consuming, full-wave solver. It is the “engine” of the ShD-based inversion and it exploits the DC block for the computation of the scattered data, $\mathbf{\Sigma}^{(v)}$, in correspondence with each coded trial solution, $\xi$;

4) Solution Space Exploration (SSE) - This block performs an effective sampling of the ISP solution space by leveraging on (a) the “hill-climbing” features of a properly customized EA strategy and on (b) the smart interaction with the DT to yield a fast and reliable convergence towards the global optimum. The SSE block receives as external inputs the samples of the incident, $\{\mathbf{T}^{(v)}(x_n, y_n), (x_n, y_n) \in D ; n = 1, ..., N\}$, and the scattered, $\{S^{(v)}(x_m, y_m), (x_m, y_m) \in O ; m = 1, ..., M\}$, fields, while it uses the unknowns coding, $\xi$, and the cost function definition, $\Phi$, from the PF block. The SSE output is the ShD solution, $\xi_{(ShD)}$, and its mapping in a contrast distribution, $\mathbf{\Sigma}^{(ShD)}$.

Each ShD block is detailed in the following by pointing out the key-item for its integrated implementation.

A. Problem Formulation (PF)

Concerning the identification of a suitable parametric model of the ISP solution in terms of a limited set of $K$ descriptors, $\xi = \{\xi_k ; k = 1, ..., K\}$, it is worth noticing that the number of DoFs $K$ is directly proportional to the size of the population of trial-solutions, $P$, used in the multiple-agent minimization of $\Phi(\xi)$, and it determines the overall computational cost of the inversion process. Therefore, it is paramount to seek for the smartest coding of the solution that minimizes the computational burden of the optimization, while enabling a careful
The extension of the spline representation to doubly-connected contours/inhomogeneous targets (e.g., piece-wise homogeneous concentric contrast distributions) as well as to multiple disconnected objects is straightforward as discussed and proved in Sect. IV.
B. Data Computation (DC)

In order to compute $\mathbf{S}^{(v)}(\xi) (v = 1, \ldots, V)$, let us remember that it is the scattered data vector radiated by the $v$-th ($v = 1, \ldots, V$) equivalent current distribution $f^{(v)}(\xi)$ according to (6). Thus, $\xi$ is first mapped into the corresponding $v$-th ($v = 1, \ldots, V$) equivalent current vector $f^{(v)}(\xi)$ whose generic $n$-th ($n = 1, \ldots, N$) entry is defined as

$$
\mathcal{J}^{(v)}(\xi) \triangleq \mathcal{T}^{(v)}(x_n, y_n|\xi) \tau(x_n, y_n|\xi).
$$

(17)

Because of the spline-based representation of the unknown scattering profile of support $\Omega$, the relation between the $n$-th ($n = 1, \ldots, N$) contrast value $\tau(x_n, y_n|\xi)$ and $\xi$ is based on the Jordan curve theorem [49] that allows one to state whether a point $(x_n, y_n)$ belongs or not to the scatterer region $\Omega$ enclosed by the spline contour $\partial \Omega(x, y|\xi)$

$$
\tau(x_n, y_n|\xi) = \begin{cases} 
\tau \Omega, & \text{if } (x_n, y_n) \in \partial \Omega(x, y|\xi) \\
0, & \text{otherwise}
\end{cases}
$$

(18)

On the other hand, the $n$-th ($n = 1, \ldots, N$) sample of the $v$-th ($v = 1, \ldots, V$) total field $\mathcal{T}^{(v)}(x_n, y_n|\xi)$ is numerically derived from the MOO-M-rectified version of (1)

$$
\mathcal{T}^{(v)}(\xi) = \left[ I - G_{D}^{T} G_{D} \right]^{-1} f^{(v)}(\xi)
$$

(19)

where $\mathcal{T}^{(v)} = \{ \mathcal{T}^{(v)}(x_n, y_n); n = 1, \ldots, N \}$, $\mathbf{f}(\xi) = \text{diag} \{ \tau(x_n, y_n|\xi); n = 1, \ldots, N \}$, $I$ is the identity matrix, and $G_{D}$ is the $(N \times N)$ internal Green’s operator whose $(n, p)$-th ($n, p = 1, \ldots, N$) entry is equal to $G_{D}^{T} = j \chi \text{J}_{D}^{T}(k_0 \rho_0) d\varepsilon d\varepsilon'$.

Once $\mathcal{T}^{(v)}(\xi) (v = 1, \ldots, V)$ has been obtained by substituting (19) and (18) in (17), the corresponding scattered field vector $\mathbf{S}^{(v)}(\xi) (v = 1, \ldots, V)$ is then computed through (6).

C. Cost Function Evaluation (CBE)

To efficiently compute the data mismatch cost function (16), by avoiding the time-consuming call to the forward (FW) solver in (5), the LBE paradigm [44] is exploited to build a fast yet accurate surrogate of $\Phi(\xi)$, $\Phi_{\text{GP}}(\xi)$, which is adaptively “reinforced” at each $i$-th ($i = 1, \ldots, I_{SD}^{\text{DT}}$) iteration of the optimization process performed in the SSE block (Sect. III-D). More specifically, a Gaussian Process (GP)-based DT [50][51] of $\Phi(\xi)$ is built at each $i$-th ($i = 1, \ldots, I_{SD}^{\text{DT}}$) iteration of the optimization, $\Phi_{\text{GP}}(\xi)$, from a training set of $S_t$ known input/output (I/O) pairs according to the following “three-step” strategy leveraging on the interconnections among all ShDoF functional blocks (Fig. 1):

- **Input-Space Reduction** - Input the minimum set of $K$ highly-informative ShDoFs (14), which univocally describe the ISP solution $\xi$, from the PF block (Sect. III-A);
- **Input-Space Representative Sampling** - Build the smallest size $i$-th ($i = 1, \ldots, I_{SD}^{\text{DT}}$) training set

$$
\Lambda_i \triangleq \left\{ \xi^{(s)}, \Phi(\xi^{(s)}); s = 1, \ldots, S_i \right\}
$$

(20)

of $S_i$ I/O pairs to suitably represent the $K$-dimensional input space. It means that for each $s$-th ($s = 1, \ldots, S_i$) sample, $\xi^{(s)}$, decoded with the DC block, $\Phi(\xi^{(s)})$ is computed with a FW solver. At the initialization ($i = 0$), the $S_t|_{i=0}$ samples are selected according to the Latin Hypercube Sampling (LHS) strategy (see Appendix I) to uniformly explore the ShDoF space thanks to its “input space filling” property [52], while new I/O pairs are adaptively selected in the SSE block and added to the training set of the previous iteration, $\Lambda_{i-1}$, to build the $i$-th ($i = 1, \ldots, I_{SD}^{\text{DT}}$) training set, $\Lambda_i$, otherwise (i.e., $1 \leq i \leq I_{SD}^{\text{DT}}$):

- **DT Generation** - Starting from the $i$-th ($i = 1, \ldots, I_{SD}^{\text{DT}}$) training set, $\Lambda_i$, define the $i$-th ($i = 1, \ldots, I_{SD}^{\text{DT}}$) GP predictor [50][51] of $\Phi(\xi)$, $\Phi_{\text{GP}}(\xi)$, as follows

$$
\hat{\Phi}_{\text{GP}}(\xi) = \chi_i + \left[ L(x_i|\xi) \right]^{T} R^{-1} \left[ L(x_i|\xi) - \chi_i \right],
$$

(21)

where $\chi_i$ is a scalar term given by

$$
\chi_i \triangleq \frac{1}{I_{SD}^{\text{DT}} - 1} \sum_{k=1}^{I_{SD}^{\text{DT}}} \Phi_{\text{GP}}(\xi)_{k,i} - \Phi(\xi)_{i,i},
$$

(22)

$\tau$ being the transpose operator, $L(x_i|\xi)$ is a $(S_i \times 1)$-dimensional vector whose $s$-th entry is equal to

$$
L(x_i|\xi)_{s,i} = \prod_{k=1}^{K} \exp \left( -\gamma_{i,k} | s^{(s)}_{k} - \xi_{k}^{(i)} |^{\beta_{i,k}} \right),
$$

(23)

$R_{k,i}$ is the $(S_i \times S_i)$ correlation matrix of $\Lambda$ whose $s$, $u$-th $(s, u = 1, \ldots, S_i)$ element is $r_{s,u}(\xi^{(s)},\xi^{(u)})$

$$
\Phi_{\text{GP}}(\xi) = \left[ \Phi(\xi^{(s)}); s = 1, \ldots, S_i \right]^{T}, \text{ and } 1_{i} \text{ is a } (S_i \times 1)
$$

unitary column vector. Moreover, $\gamma_{i,k}$ and $\beta_{i,k}$ are the $k$-th ($k = 1, \ldots, K$) elements of the GP hyper-parameter vectors $\gamma$ and $\beta$, respectively, which are yielded from the maximization of the concentrated log-likelihood function [50]

$$
\Gamma(\gamma, \beta) = -\frac{1}{2} \left\{ S_t \times \ln \left( \nu_t \right) + \ln \left( \det \left( R \right) \right) \right\},
$$

(24)

where

$$
\nu_t \triangleq \frac{1}{S_t} \left[ \left( \Phi_{\text{GP}}(\xi) - 1 \chi_i \right)^{T} R_{i,i}^{-1} \left( \Phi_{\text{GP}}(\xi) - 1 \chi_i \right) \right],
$$

(25)

$\ln(\cdot)$ and $\det(\cdot)$ being the natural logarithm and the matrix determinant operators.

It is worth noticing that the choice of the GP to build the DT of $\Phi(\xi)$, unlike other regression strategies such as, for instance, the Support Vector Regression (SVR) [44], ensures an exact prediction of the actual value of the cost function when a trial solution, $\xi$, coincides with a training sample, $\xi^{(s)}$ (i.e., $\hat{\Phi}_{\text{GP}}(\xi^{(s)}) = \Phi(\xi^{(s)}); s = 1, \ldots, S_t$). Moreover, it must be pointed out that the definition of the $i$-th ($i = 1, \ldots, I_{SD}^{\text{DT}}$) GP surrogate model, $\Phi_{\text{GP}}(\xi)$ in (21) is based on the assumption that the actual value of the cost function, $\Phi(\xi)$, is the realization of a normally-distributed random variable with average value $\Phi_{\text{GP}}(\xi)$ and variance $[50]$ equal to

$$
\delta^2(\xi) = \nu_t^2 \left[ 1 - \left| L(x_i|\xi) \right|^{T} R_{i,i}^{-1} \left| L(x_i|\xi) \right| + \frac{1 - L(x_i|\xi)^{T} R_{i,i}^{-1} L(x_i|\xi)}{1} \right],
$$

(26)
This latter quantity provides an estimate of the reliability of the GP-based DT, greater values of $\delta_i^2(\xi)$ corresponding to a lower “reliability” of the associated prediction $\hat{\Phi}_i(\xi)$. Indeed, the value of $\delta_i^2(\xi)$ depends on $\xi_i(\xi)$ (26), which in turn is related to the $\xi_{(s)}$-weighted distance between $\xi_i$ and the $s$-th ($s = 1, ..., S_i$) training sample, $\xi_{(s)}$ (23). Thus, if $\xi_i$ is very far from all the $S_i$ training samples, $\xi_{(s)}$, then $\xi_i(\xi) \to 0$ and the uncertainty reaches its maximum (i.e., $\delta_i^2(\xi) \to \nu_i^2$). On the contrary, the uncertainty is minimal in correspondence of the training samples [i.e., $\delta_i^2(\xi_{(s)}) = 0$ ($s = 1, ..., S_i$)]. Finally, let us consider that, according to the GP theory [51], the actual value of the cost function $\Phi(\xi)$ fulfills at least to 95% probability [51] the following condition

$$L_i(\xi) \leq \Phi(\xi) \leq U_i(\xi),$$

$L_i(\xi)$ and $U_i(\xi)$ being the lower and the upper “confidence bounds”, respectively, defined as

$$L_i(\xi) = \hat{\xi}_i(\xi) - 2\delta_i(\xi),$$
$$U_i(\xi) = \hat{\xi}_i(\xi) + 2\delta_i(\xi)$$

so that $L_i(\xi_{(s)}) = U_i(\xi_{(s)}) = \hat{\Phi}_i(\xi_{(s)}) = \Phi(\xi_{(s)})$ ($s = 1, ..., S_i$).

D. Solution Space Exploration (SSE)

To explore in a smart way the $K$-dimensional ShD solution space for solving the non-linear ISP, nature-inspired EAs are the most suitable candidates to effectively implement such a task without requiring, unlike DOs, the differentiation of the data mismatch cost function (16) [32]. However, a “bare” integration of an EA-GO with a forward solver (FW) would imply an overall inversion time equal to

$$\Delta t_{GO} = (P \times I_{GO}) \times \Delta t_{FW},$$

$P$ and $\Delta t_{FW}$ being the number of trial solutions evolved through $I_{GO}$ iterations and the time of a single full-wave evaluation of (16), which clearly becomes unpractical in many applicable scenarios requiring a fast inversion. If a significant reduction of $P$ can be yielded with a minimum-dimensionality coding of the unknown scattering profiles (e.g., the spline-based strategy in Sect. III-A), it is not enough towards a computationally-competitive global inversion/optimization. In order to break down the computational burden required by the iterated (multi-agent) evaluation of (16) to comply with (7) by reducing $\Delta t_{GO}$ (29), there are two different strategies. The former is that of minimizing the number of iterations of the EA to reach the global optimum $\xi_{(opt)}$, $I_{GO}$. Towards this end, it is mandatory to choose an EA that provides a proper balance between exploration and exploitation to enable “hill-climbing” features for effectively escaping from local minima/false solutions as well as to guarantee a quick convergence towards the attraction basin of the global minimum of the cost function $\Phi(\xi)$. Accordingly, the Particle Swarm Optimization (PSO) algorithm [32] is chosen as a robust and effective evolutionary strategy particularly suitable for the exploration of the real-valued solution space of the ShD-DoFs (14). During $I_{ShD}$ iterations, the PSO processes a swarm of $P$ particles/agents, $A = \{ A_p; p = 1, ..., P \}$, by changing their velocities, $V = \{ \xi_p; p = 1, ..., P \}$, to evolve their positions in the solution space, $P = \{ \xi_p; p = 1, ..., P \}$, until reaching the global optimum (i.e., $\xi_{(opt)} = \arg \min \{ \Phi(\xi) \}$).

The second method to shorten (29) is that of building a surrogate model in the CFE block (see Sect. III-C) to replace the FW solver during the optimization so that $\Delta t_{test} \ll \Delta t_{FW}$. However, the definition of a globally-accurate predictor would generally require a huge number of training samples $S$ [ $S \gg (P \times I)$], which not linearly depends on the number of scatterer descriptors, $K$, because of the so-called “curse-of-dimensionality” [43]. On the other hand, it is worth to consider that the DT is required to predict the value of the cost function $\Phi(\xi)$ (16) for guiding the GO search throughout the solution space with an accuracy adaptively enhanced and very high only in the attraction basin (i.e., in the proximity) of the global optimum. Owing to such considerations, a “collaborative” framework is implemented between the PSO, which is responsible of sampling the solution space with the swarm $A$ of $P$ trial agents, and the DT model based on the GP regression strategy [50][51] that gives not only a prediction of the cost function associated to each trial solution, $\hat{\Phi}(\xi_p)$ ($p = 1, ..., P$), but also an estimate of its “degree
of reliability”, \( \delta(\xi^{(p)}) \). This latter is an additional information to be profitably exploited for identifying “promising” solutions for which the cost function (16) is expected to be lower than any previously-explored solution set. Moreover, the value \( \delta \) can be used as a threshold for triggering adaptive refinements/reinforcements, obtained by simulating selected particles to enhance the accuracy only “where needed”, of the predictor during the optimization loop. The resulting \( \text{SSE block} \) then works as follows:

1) **Initialization (i = 0)** - With the CFE block (Sect. III-C), build the initial training set of \( S_0 \) I/O pairs, \( \Lambda_0 = \left\{ \left( \xi(s), \Phi(\xi^{(s)}) \right) : s = 1, ..., S_0 \right\} \), to train the initial GP predictor \( \Phi_0(\xi) \). Randomly initialize the positions of the swarm \( \Lambda_0 \) of \( P \) particles, \( \Psi_0 = \left\{ \xi^{(p)} : p = 1, ..., P \right\} \), with random velocities, \( \xi_0 = \left\{ \xi_k^{(p)} : p = 1, ..., P \right\} \), and set the personal best position of each \( p \)-th particle (\( p = 1, ..., P \)) to the initial one (i.e., \( \xi^{(p)} = \xi_k^{(p)} \));

2) **SbD Optimization Loop (i = 1, ..., I_{SbD})**
   a) **Cost Function Prediction** - For each \( p \)-th \( (p = 1, ..., P) \) particle of the current \( i \)-th swarm, \( \Lambda_i \), predict the values of \( \Phi \left( \xi^{(p)} \right), \mathcal{L} \left( \xi^{(p)} \right) \), and \( \Upsilon \left( \xi^{(p)} \right) \), with the \( i \)-th DT \( \Phi_i(\xi) \);
   b) **Particles Ranking** - Determine the “best promising” \( (BP) \) position of a particle of \( \Lambda_i \)
   \[
   \xi^{(BP)} = \arg \left\{ \min_{p=1,..,P} \left[ \mathcal{L}_i \left( \xi^{(p)} \right) \right] \right\} ; \quad (30)
   \]
   c) **DT Adaptive Updating** - If \( \mathcal{L} \left( \xi_i^{(BP)} \right) < \min_{s=1,..,S_i} \left[ \Phi \left( \xi^{(s)} \right) \right] \) perform the following operations, otherwise set \( \hat{S}_i \leftarrow \hat{S}_{i-1} \) and \( \Lambda_i \leftarrow \Lambda_{i-1} \) and jump to Step 2(d):
   
   i) Exploit the DC bock (Sect. III-B) to derive the \( v \)-th \( (v = 1, ..., V) \) induced equivalent current, \( \mathcal{G}^{(v)} \left( \xi_i^{(BP)} \right) \) from \( \xi_i^{(BP)} \), then compute the corresponding scattered field, \( \xi^{(BP)} \left( \xi_i^{(BP)} \right) \);
   ii) Compute \( \Phi \left( \xi_i^{(BP)} \right) \) with (16);
   iii) Update the training set by adding the BP training set sample, \( \Lambda_i \leftarrow \Lambda_{i-1} \cup \left\{ \xi^{(BP)} \left( \xi_i^{(BP)} \right) \right\} \), and let \( \hat{S}_i \leftarrow \left( \hat{S}_{i-1} + 1 \right) \);
   iv) Use the CFE block (Sect. III-C) to re-train the GP predictor using the updated/reinforced training information within \( \Lambda_i \);

d) **Personal Best Updating** - Update the personal best position of each \( p \)-th \( (p = 1, ..., P) \) particle, \( \hat{\psi}_{i,k} = \hat{\psi}_{i,k} \left( \psi_{i,k} : k = 1, ..., K \right) \), according to the \( \text{SbD-updating rules in Fig. 3(a);} \)

e) **Global Best Updating** - Update the global best, \( \hat{\psi} = \left\{ \hat{\psi}_{i,k} : k = 1, ..., K \right\} \) according to the workflow in Fig. 3(b);

f) **Convergence Check** - Stop the optimization if \( i = I_{SbD} \) and output the SbD solution set, to the current global best swarm position, \( \hat{\xi}_{SbD} = \hat{\psi} = I_{SbD} \left( x_n, y_n \right) : n = 1, ..., N \) yielded from the DC bock (Fig. 1 - Sect. III-B). Otherwise, proceed to Step 2(g);

g) **Velocities Updating** - Update the velocity vector \( \hat{V}_i \rightarrow \hat{V}_{i+1} \) by computing the \( k \)-th \( (k = 1, ..., K) \) component of the velocity of the \( p \)-th \( (p = 1, ..., P) \) particle of the swarm \( \Lambda_{i+1} \) according to the \( \text{PSO} \) mechanism
   \[
   \xi_{i+1,k}^{(p)} = w \xi_{i,k}^{(p)} + \ell_1 \xi_{i,k}^{(p)} - \xi_{i,k}^{(p)} + \ell_2 \xi_{i,k}^{(p)} - \psi_{i,k} \]
   \[
   + \ell_2 \xi_{i,k}^{(p)} - \psi_{i,k} \]
   \[
   \]
This section is aimed at presenting a set of representative numerical and experimental results drawn from an extensive validation of the proposed SbD-based inversion method. Unless stated otherwise, a square investigation domain D of side \( L_D = 2 \times \lambda \) has been probed by \( V = 18 \) incident plane waves impinging from the \( V \) angular directions \( \varphi_v \) of the form \( \varphi_v = 2\pi \frac{(v-1)}{V} \) \( v = 1, ..., V \). The scattered field samples have been collected at \( M = 18 \) probing locations uniformly distributed on a circular observation domain \( O \) of radius \( r_0 = 3 \times \lambda \). As for the generation of the synthetic scattered field data, the MoM solution of the FW problem (1)(5) has been performed by partitioning the investigation domain into \( N_{FW} = 40 \times 40 \) square sub-domains, while \( N = 20 \times 20 \) pixel bases have been adopted in the inversion process to avoid the inverse crime after applying the pixel-mapping operator (18) in the DC Block (see [1] p. 174). Moreover, an additive Gaussian noise has been added to the synthetically-generated data samples to test the robustness of the inversion to different signal-to-noise ratios (SNRs). Furthermore, owing to the stochastic nature of the SbD-based approach, a set of \( T = 50 \) random executions has been run for each inversion dataset to ensure the statistic meaningfulness of the results.

Concerning the imaging results/performance and besides the pictorial representation of the reconstruction in terms of color-

Figure 6. Numerical Assessment (Test Case #1): \( \tau_0 = 4.0, V = M = 18 \), Noiseless Data; \( K = 8 \) - Contrast profile associated to 4 training samples.

Figure 7. Numerical Assessment (Test Case #1): \( \tau_0 = 4.0, V = M = 18 \), Noiseless Data; \( K = 8 \) - Evolution of the optimal value of the cost function, \( \Phi_{\text{i}} \), versus the iteration index, \( i \).

h) Swarms Updating - Update the position vector (\( P_i \rightarrow P_{i+1} \)) by adding to the \( k \)-th component of the current position of the \( p \)-th particle of the swarm \( A_{i+1} \) the corresponding term of the velocity vector \( V_{i+1} \)

\[
\xi_{i,k+1}^{(p)} = \xi_{i,k}^{(p)} + \eta_{i,k+1}^{(p)}, \tag{32}
\]

then let \( i \leftarrow (i + 1) \) and go to Step 2(a).

It is worth pointing out that the SSE block implements a novel “time-constrained reinforced PSO” strategy to allow the user to \textit{a-priori} fulfill the CPU-time target (7) by properly setting the size \( S_0 \) of the initial training set, \( \Delta_0 \), and the maximum number of \( DT \) “reinforcements”, \( I_{SbD} \), performed during the global minimization of (16). Indeed, the total number of calls to the FW solver during a SbD inversion, thus the SbD time cost, as well, is upper-bounded to \( S = (S_0 + I_{SbD}) \) so that a SbD inversion turns out to be computationally advantageous with respect to a standard GO solution when \( S \ll (P \times I_{GO}) \), with a time saving index equal to

\[
3 \text{If } S \ll 10^3 \text{ (Sect. IV), the overall time required to train } (\Delta t_{\text{train}}^{(p)}) \text{ and to test } (\Delta t_{\text{DT}}^{(p)}) \text{ the DT model can be neglected since } I_{SbD} \times \Delta t_{\text{train}}^{(p)} \ll \Delta t_{FW} \text{ and } P \times I_{SbD} \times \Delta t_{\text{DT}}^{(p)} \ll \Delta t_{FW} \text{ [43].}
\]
maps of the dielectric profile of $D$, the accuracy of the data inversion is quantified by the error index

$$
\Xi = \frac{1}{N} \sum_{n=1}^{N} \frac{|\tau(x_n, y_n) - \bar{\tau}(x_n, y_n)|}{|\tau(x_n, y_n) + 1|},
$$

(34)

where $\tau(x_n, y_n)$ and $\bar{\tau}(x_n, y_n)$ stand for the actual and the retrieved complex-valued contrast value of the $n$-th ($n = 1, ..., N$) pixel $D_n$ ($D_n \in D$), respectively.

The first test case deals with the noiseless reconstruction of the scattering profile in Fig. 4(a) having contrast $\tau_D = 4.0$. The $SbD$-based inversion has been carried out by considering a spline description of the scatter with $Q = 4$ control points ($\Rightarrow K = 8$ - Tab. I) and choosing, according to the guidelines in [32] a swarm size of $P = 10$ particles, a constant inertial weight equal to $w = 0.4$, and acceleration coefficients with values $\bar{\ell}_1 = \bar{\ell}_2 = 2.0$. To investigate on the dependence of the prediction accuracy of the DT inversion method that exploits the same spline-based coding, to a "bare" $\Phi_1$, versus the iteration index, $i$, and of (b) the reconstruction error, $\Xi$, and the execution time, $\Delta t$, versus the SNR value of the scattered data.

In a less evident way, the ratio $S_0/K = 5$ ($\rightarrow |\eta|_{S_0/K=5} \approx 7\%$ and $\Delta t_{\text{sav}}|_{S_0/K=5} = 86\%$) has been then chosen as the optimal trade-off threshold to fit (7). More specifically, the size of the initial and the final training datasets have been set here to $S_0 = 5 \times K = 40$ and $S = 140$, respectively, so that the total execution time of the $SbD$ is equal to that of a DO method (i.e., $\Delta t_{\text{SD}} \approx \Delta t_{\text{DO}}$), which is based on a standard implementation of the Conjugate Gradient (CG) technique, running for $I_{\text{DO}} = 400$ iterations [24]. For completeness, the contrast distributions of 4 training samples is shown in Fig. 6, verifying that the training set comprises a large variability of objects in terms of shape, position, and dielectric properties, sharing with the actual object only the (a-priori chosen) number of control points, $Q$. By using such a setup, Figure 7 shows the evolution of the optimal value of the cost function, $\Phi$, versus the iteration index, $i$, and of (b) the reconstruction error, $\Xi$, and the execution time, $\Delta t$, versus the SNR value of the scattered data.

5A standard implementation of the CG has been considered, exploiting a null contrast as initial guess, the maximum number of iterations as stopping criterion, and the presence of the "state term" in the cost function as regularization strategy [8].
the solution space looking for the global optimum even though guided by a DT model of the FW solver. For completeness, the behavior of the DO minimization is reported, as well. To better understand the optimization performance of the three inversion approaches, Figure 8 shows the 2-D parametric representation of the functional described by the following equation

$$
\Phi (a,b) = \Phi \left\{ b \times \left[ (a+1) \times \xi (1) - a \times \xi (act) \right] + \right\} + (b-1) \times a \times \xi (2) \right\}
$$

in the ranges $-1.5 \leq a \leq 0.5$ and $-0.5 \leq b \leq 1.5$ when setting $\xi (1) = \xi (SbD)$ and $\xi (2) = \xi (DO)$ for $\xi (2) = \xi (Go)$ [Fig. 8(b)], $\xi (act)$ being the actual solution. The landscape in Fig. 8(a) proves that the DO solution is trapped into a local minimum of the cost function (i.e., a false solution for the inversion) without any possibility to escape from such a “wrong” valley. This is even more evident by looking at the plot of the cost function along the $1-D$ cut of the solution space passing through $\xi (DO)$ and the actual solution [i.e., $\phi (a,b) \xi (2) = \xi (DO)$ - Fig. 8(c)]. Otherwise, the SbD solution $\xi (SbD)$ belongs to the “attraction basin” of the actual solution $\xi (act)$ analogous to the GO solution $\xi (GO)$ [Figs. 8(b)-8(c)]. Such outcomes are confirmed by the corresponding reconstructions in Figs. 4(b)-4(d). Indeed, the DO inversion is unsatisfactory and remarkably worse than the SbD one, as quantified by the integral errors (i.e., $\xi_{SNR} = 51.3$ - Fig. 9), even though the execution time of the two iterative minimizations is approximately the same ($\Delta t_{DO} = 480$ [sec] vs. $\Delta t_{SbD} = 490$ [sec] - Fig. 9). Furthermore, the computational efficiency of the SbD is disruptive when compared to the standard GO since $\Delta t_{SbD}^{max} = 0.14$ ($\Rightarrow \Delta t_{SbD} = 86%$ - Fig. 9), while yielding the same accuracy (i.e., $\xi_{SNR} = 0.99$ - Fig. 9).

In order to assess the robustness of the reconstruction process to blurred/corrupted data, Figure 10(a) compares the behavior of the cost function for the SbD and the GO optimizations when varying the $SNR$ of the scattered field samples. As expected, the data matching gets worse as the noise increases from $SNR = 20$ [dB] up to $SNR = 5$ [dB] (i.e., $\Phi \xi (SbD) \mid SNR=20$ [dB] = $6.84 \times 10^{-2}$ vs. $\Phi \xi (SbD) \mid SNR=10$ [dB] = $1.32 \times 10^{-1}$ vs. $\Phi \xi (SbD) \mid SNR=5$ [dB] = $2.70 \times 10^{-1}$ - Fig. 10(a)), but the SbD still performs as the GO, while reducing the inversion time $\Delta t_{SbD} = 86%$ - Fig. 10(b)), despite the need of predicting the cost function values starting from non-ideal (blurred) data. The reliability of the SbD in emulating a GO when exploring highly-nonlinear solution spaces is confirmed by the comparison of the corresponding reconstruction errors

$\xi (SbD) = \xi (Go)$ for $\xi (act)$. The $x_i$ is worth pointing out that a standard definition of the DO-DOs, i.e., $\xi (Do) = \{\xi^{(i)} (x_i, y_i) \times \xi (act) \times \xi (Do) \times \xi ^{(-1)(0)} ; v = 1, \ldots, V ; n = 1, \ldots, N \}$, has been adopted according to the reference literature on gradient-based local search algorithms [8].

According to (36), it can be easily verified that $\Phi (-1,1) = \Phi \xi (act)$.
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAP.2021.3139627, IEEE Transactions on Antennas and Propagation.

Figure 13. Numerical Assessment (Test Case #1): $V = M = 18$, $SNR = 5$ [dB]; $K = 8$ - Plot of (a) the evolution of the optimal value of the cost function, $\Phi_\Omega$, versus the iteration index, $i$, and of (b) the reconstruction error, $\Xi$, and the execution time, $\Delta t$, versus the value of the contrast of the scatterer, $\tau_\Omega$.

Figure 14. Numerical Assessment (Test Case #1): $V = M = 18$, $SNR = 5$ [dB]; $K = 8$ - Reconstructions of the contrast profile in $D$ obtained by (a)-(c) the $SbD$, (d)-(f) the $GO$, and (g)-(i) the $DO$ when the actual value of the contrast of the scatterer is (a)/(d)/(g) $\tau_\Omega = 1$, (b)/(e)/(h) $\tau_\Omega = 2$, and (c)/(f)/(i) $\tau_\Omega = 10$.

Figure 15. Numerical Assessment (Test Case #2): $\tau_\Omega = 4.0$, $V = M = 18$, $SNR = 10$ [dB]; $K = 12$ - Plot of (a) the evolution of the optimal value of the cost function, $\Phi_\Omega$, versus the iteration index, $i$, and, color maps of the functional (36) (c)/(d) in the ranges $-1.5 \leq a \leq 0.5$ and $-0.5 \leq b \leq 1.5$, when setting $\varepsilon^{(1)} = \varepsilon^{(SbD)}$ and (c) $\varepsilon^{(2)} = \varepsilon^{(GO)}$ or (d) $\varepsilon^{(2)} = \varepsilon^{(DO)}$. 

[Fig. 10(b)], which are almost identical whatever the amount of noise and, always, significantly lower than the DO ones. As a matter of fact, the DO is unable either to find a satisfactory reconstruction [Figs. 12(g)-(i)] or to localize the attraction basin of the global optimum [Fig. 11(a), Fig. 11(c), and Fig. 11(e)]. It is also worth noticing that the $SbD$ is effective even under very harsh operative conditions (e.g., $SNR = 5$ [dB]) as confirmed pictorially in Fig. 12(c) and quantitatively by the value of the error index [i.e., $\Xi_{SbD|SNR=5\,\text{[dB]}} = 5.6 \times 10^{-2}$ - Fig. 10(b)].

The next set of results are concerned with the dependence of the data inversion on the contrast value of the scatterer, $\tau_\Omega$, still considering the extremely challenging scattering environment with $SNR = 5$ [dB]. Figure 13(a) gives some indications on the iterative minimization of the cost function. As expected, the weaker the scatterer more effective is the optimization process as denoted by the smaller and smaller values of the cost function at the convergence [i.e., $\Phi_\Omega(\varepsilon^{(SbD)})|_{\tau_\Omega=1} = 2.54 \times 10^{-1}$, $\Phi_\Omega(\varepsilon^{(SbD)})|_{\tau_\Omega=2} = 2.57 \times 10^{-1}$, and $\Phi_\Omega(\varepsilon^{(SbD)})|_{\tau_\Omega=10} = 5.72 \times 10^{-1}$ being $\Phi_\Omega(\varepsilon^{(GO)})|_{\tau_\Omega=1} = 2.50 \times 10^{-1}$, $\Phi_\Omega(\varepsilon^{(GO)})|_{\tau_\Omega=2} = 2.56 \times 10^{-1}$, and $\Phi_\Omega(\varepsilon^{(GO)})|_{\tau_\Omega=10} = 5.56 \times 10^{-1}$]. This implies that the reconstruction quality decreases as $\tau_\Omega$ increases [Fig. 13(a) and Fig. 14]. However, it has to be observed that the performance of the $GO$-based methods are significantly better...
than those from the DO, which results unable to handle high contrasts [e.g., $\tau_\Omega = 10$ - Fig. 14(i)] that cause high nonlinearities.

The second test case is related to a more complex scatterer profile [Fig. 17(a)] with $\tau_\Omega = 4.0$ and described by a larger number of spline control points ($Q = 8$), thus a greater dimensionality (i.e., $K = 12$ - Tab. I) of the solution space. Therefore, a larger initial training set has been chosen to keep the optimal setup of the $S_0/K$ ratio (i.e., $S_0 = 5 \times K = 60$), while the number of $SbD$ iterations has been reduced (i.e., $I_{SbD} = I_{GO} = 80$) to fit the time constraint (7). Despite the smaller number of optimization iterations, the higher dimensionality, and the non-negligible noise level of the scattered data (i.e., $SNR = 10$ [dB]), the $SbD$ solution is very close to the $GO$ one [i.e., $\Phi(\xi^{(SbD)}) = 2.07 \times 10^{-1}$ vs. $\Phi(\xi^{(GO)}) = 1.79 \times 10^{-1}$ - Fig. 15(a)] and, unlike the DO, it belongs to the "attraction basin" of the actual solution [Figs. 15(b)-15(c)]. Consequently, the retrieved contrast distributions [Figs. 17(b)-17(d)] quite faithfully reproduce the actual one [Fig. 17(a)] with similar values of the reconstruction error and significantly smaller than those of the DO (i.e., $\Xi|_{SbD} = 5.46 \times 10^{-2}$ vs. $\Xi|_{GO} = 4.54 \times 10^{-2}$ vs. $\Xi|_{DO} = 4.36 \times 10^{-1}$ - Fig. 16). On the other hand, the CPU-time of the $SbD$ inversion is remarkably lower than that of the GO (i.e., $\Delta t_{sav} = 82.5\%$ - Fig. 16) and very close to the DO.

Clearly, a-priori information on the unknown profile helps in selecting the most suitable number of spline control points, $Q$, used to perform the inversion. However, Fig. 18 reports the outcomes of the $SbD$ when retrieving the same target using a lower [$Q = 3$ - Fig. 18(a)] or higher [$Q = 12$ - Fig. 18(b)] number of control points with respect to the actual one [i.e., $Q = 8$ - Fig. 17(a)]. Interestingly, the object can still be correctly located (although its contrast seems over-estimated) even when retrieving it with an insufficient number of DoFs to model its complex-shaped contour [i.e., $Q = 3$ - Fig. 18(a)]. On the other hand, it can be inferred that a larger number of points always guarantees an accurate solution, as verified by the plot in Fig. 18(b) and by looking at the behavior of the reconstruction error in function of $Q$, remarking the robustness of the proposed methodology [Fig. 18(c)].

The feasibility of representing doubly connected (DC) contours/inhomogeneous objects is assessed in the third test case [Fig. 20(a)]. More in detail, the scatterer has been modeled with the following set of $K = 11$ (i.e., $S_0 = 5 \times K = 55$, $I_{SbD} = I_{GO} = 85$) descriptors

$$\xi_{DC} = \left\{ x_{\Omega}, y_{\Omega}, R\left(\xi^{(out)}_{\Omega}\right), \Im\left(\xi^{(out)}_{\Omega}\right), \Re\left(\xi^{(int)}_{\Omega}\right), \Im\left(\xi^{(int)}_{\Omega}\right), P^{(out)}; v \right\} \quad (37)$$

where the superscript $\left\{ \text{out} \right\}$ $\left\{ \text{int} \right\}$ refers to the outer [internal] contour $\partial\Omega^{(\text{out})}$ $\partial\Omega^{(\text{int})}$, while $0 < v < 1$ is the scale factor.
Table 1

| Object               | $K$ | $S_0$ | $\Delta t^{(\text{int})}$ [sec] | $\Delta t^{(\text{out})}$ [sec] | $I_{SBD}$ | Profile   | $\rho$ [\lambda] | $x, y$ [\lambda] | $Q$ | $\gamma$ [\lambda] |
|----------------------|-----|-------|-------------------------------|---------------------------------|----------|-----------|------------------|------------------|-----|-------------------|
| Fig. 4(a)            | 8   | 40    | $1.80 \times 10^{-5}$          | $1.1 \times 10^{-5}$            | 100      | $\partial t$ | (0, 0)           | 4                | {0.5, 0.5, 0.5, 0.5} |
| Fig. 6(a)            | 12  | 60    | $1.90 \times 10^{-5}$          | $1.2 \times 10^{-5}$            | 80       | $\partial t$ | (0.1, 0.1)       | 8                | {0.6, 0.2, 0.2, 0.4, 0.6, 0.6, 0.1} |
| Fig. 18(a)           | 11  | 55    | $1.90 \times 10^{-5}$          | $1.2 \times 10^{-5}$            | 85       | $\partial t^{(\text{int})}$ | (0.2, 0.2)       | 4                | {0.6, 0.6, 0.6, 0.6} |
| Fig. 18(b)           | 11  | 55    | $1.90 \times 10^{-5}$          | $1.2 \times 10^{-5}$            | 85       | $\partial t^{(\text{out})}$ | (0.2, 0.2)       | 4                | {0.36, 0.36, 0.36} |
| Fig. 20(a)           | 16  | 80    | $2.05 \times 10^{-5}$          | $1.3 \times 10^{-5}$            | 60       | $\partial t^{(1)}$ | (0.4, 0.4)       | 4                | {0.4, 0.4, 0.4, 0.4} |

Figure 19. **Numerical Assessment (Test Case #3): $V = M = 18$, $SNR = 10$ [dB]; $K = 11$** - Values of the reconstruction error, $\Xi$, and total inversion time, $\Delta t$, for the scattering scenario in Fig. 18(a) when $(a)$ ($t^{(\text{out})}_1 = 3$, $t^{(\text{init})}_1 = 0$) and $(b)$ ($t^{(\text{out})}_1 = 2$, $t^{(\text{init})}_1 = 4$).

between the two borders, the $q$-th ($q = 1, ..., Q; Q = 4$) control point of $\partial t^{(\text{init})}$ [i.e., $\rho^{(\text{init})} = \{\rho[q, \text{int}]; q = 1, ..., Q\}$] being $\rho^{(\text{q, int})} = y_p [\Omega_q]$ (Tab. I). The outcomes from such a benchmark are summarized in Fig. 19(a) in terms of reconstruction errors and execution time. Once again, these results confirm the superior trade-off between computational efficiency and effectiveness of the SBD method over the GO and the DO ones. As for the retrieved contrast, Fig. 20 shows that the SBD reconstruction provides a reliable estimation of both the object shape and the contrast value ($t^{(\text{out})}_1 = 3$, $t^{(\text{init})}_1 = 1$, $v = 0.6$ - Tab. I) well detecting the presence of a “hole” $[\Xi_{SBD} = 3.30 \times 10^{-2}$ - Fig. 20(c) vs. Fig. 20(a)]. Similar outcomes can be drawn [Fig. 19(b)] for the inhomogeneous profile in Fig. 20(b) ($t^{(\text{out})}_1 = 2$, $t^{(\text{init})}_1 = 4$, $v = 0.4$ - Tab. I), the dielectric profile inferred by the SBD being shown in Fig. 20(d) $[\Xi_{SBD} = 5.13 \times 10^{-2}$ - Fig. 19(b)].

The extension to multiple objects (MO) is dealt with in the Test Case #4 where two disconnected scatterers have been considered. In this case, the $K = 16$ unknowns are

$$\Xi_{MO} = \left\{ x^{(1)}, y^{(1)}, \ldots, y^{(4)}; x^{(2)}, y^{(2)}; \Re \left( \frac{\tau^{(1)}}{\lambda} \right), \Im \left( \frac{\tau^{(1)}}{\lambda} \right), \Re \left( \frac{\tau^{(2)}}{\lambda} \right), \Im \left( \frac{\tau^{(2)}}{\lambda} \right) \right\},$$

where the superscripts $(1)/(2)$ refer to the two disconnected spline contours $\partial t^{(1)}/\partial t^{(2)}$ ($Q = 4$), and the SBD has been run for $I_{SBD} = I_{GO} = 60$ iterations starting from a training set with $S_0 = 5 \times K = 80$ I/O pairs. Despite the higher complexity of the ISP problem at hand, also related to a larger dimension of the solution space as well as the non-negligible contrast of both scatterers ($\tau^{(1)} = \tau^{(2)} = 4$), the SBD carefully images the investigation domain [i.e., $\Xi_{SBD} = 1.05$ - Fig. 22(b) vs. Fig. 22(c) and $\Xi_{SBD} = 8.9 \times 10^{-2}$ - Fig. 22(b) vs. Fig. 22(d)] by reducing the inversion time of about $\Delta t_{SBD} = 76.7\%$ (Fig. 21).

Finally (Test Case #5), the SBD-based imaging method has been assessed against laboratory-controlled experimental data.
Numerical Assessment (Test Case #4: $\Omega^{(1)} = \tau^{(2)} \Omega = 4, V = M = 18, \text{SNR} = 10 \text{[dB]}$; $K = 16$) - Values of the reconstruction error, $\Xi$, and total inversion time, $\Delta t$.

![Figure 21](image1)

Experimental Assessment (Test Case #5: $f = 2 \text{[GHz]}$, $\tau^{\text{(out)}} = 0.45, \tau^{\text{(int)}} = 2, V = 8, M = 241; K = 11$) - Maps of (a) the actual “FoamDielInt” [53] and (b)-(d) the retrieved contrast distributions with (b) the SbD, (c) the GO, and (d) the DO methods.

![Figure 23](image2)

With reference to the data provided by the Institut Frésnel [53], the “FoamDielInt” scattering scenario has been selected as representative benchmark. It consists of a foam cylinder with diameter $8.0 \times 10^{-2}$ [m] and contrast $\tau^{\text{(out)}} = 0.45$ that embeds a smaller, $3.1 \times 10^{-2}$ [m] in diameter, and weaker, $\tau^{\text{(int)}} = 2.0$, dielectric cylinder [Fig. 23(a)]. The acquisition system was composed by $V = 8$ ridged-horn antennas working at $f = 2$ [GHz] to probe a square investigation domain $D$ of side $L_D = 0.2$ [m]. The scattered data have been collected in $M = 241$ uniformly-spaced locations on a circular observation domain $O$ with radius $r_O = 1.67$ [m] [53]. Because of the topology of the object at hand, the exploration of the solution space defined by the DoFs in (37) has been carried out by letting $S_0 = 55$ and $I_{SbD} = I_{GO} = 85$ according to the previous examples. Figure 23(b) shows the retrieved contrast distribution. Similarly to the GO image [Fig. 23(c)], it is possible to detect the two-layers scatterer with a reliable estimation of the outer support of the object, $\partial \Omega^{\text{(out)}}$, as well as to infer the presence of an inner scatterer/layer with higher permittivity. Once again, it turns out that it is possible to address the problem of local minima by exploiting the “hill-climbing” features of an EA-based multiple-agent approach, but solving the arising global minimization task with a remarkable time saving over a standard GO implementation (i.e., $\Delta t_{\text{GO}} = 83.5\%$) by equalling the computational efficiency of the DO [Fig. 23(d)].

V. CONCLUSIONS

An innovative strategy has been proposed to address the computationally-efficient yet reliable solution of the fully non-linear ISP. The inversion method has been built by implementing the pillar concepts of the SbD framework [43] to allow an effective exploration of the multi-modal landscape defined by the data mismatch cost function with the same time cost of a standard deterministic local search.

From a methodological point of view and to the best of the authors’ knowledge, the key advances of this research work with respect to the state-of-the-art literature can be summarized as follows:

- a “smart” and flexible minimum-dimensionality encoding of complex-shaped scatterers yielded with a spline-based modeling of the scattering profile (Sect. III-A), which not only “implements” a more favorable “operating environment” for the underlying EA-based GO strategy, but it also alleviates the “curse-of-dimensionality” problem;
- the use of a GP-based LBE approach for building a fast and accurate DT of the time-consuming FW solver that predicts the data mismatch cost function associated to each trial solution, but also provides additional information on the associated “confidence level” of this latter;
- the setup of a collaborative framework between the EA mechanisms and the DT model that enables an effective
exploration of the solution space, which is adaptively sampled at selected and promising points to increase the prediction accuracy of the DT model as well as to speed-up the converge towards the attraction basin of the global-optimum-actual-solution.

Moreover, the main outcomes from the numerical and experimental assessment (Sect. IV) are:

- the Shb-based inversion method is a reliable tool for reaching the attraction basin of the global optimum without being trapped into local-minima/false-solutions also when highly nonlinear cost functions/strong scatterers are at hand;
- it exhibits the same computational efficiency of a DO, breaking - for the first time to the authors' best knowledge - the widely-diffused idea that solving an ISP with an EA-based tool is generally computationally unaffordable;
- the range of a reliable and effective application of the Shb inversion method extends from weak to strong simple as well as complex and multiple/inhomogeneous objects in harsh environmental conditions, as well, subject to a suitable choice of the Shb building blocks according to the "no-free lunch" theorems [54];
- the Shb inversion is able to effectively and efficiently process synthetic as well as real laboratory-controlled scattering data.

As for the DO considered for the comparisons, it is worth remarking that the CG has been selected as a representative example of a standard IS solution approach not benefiting from the problem-driven selection, customization, and integration of the Shb functional blocks enabling the effective/efficient solution of the problem at hand.

Future works, beyond the scope of this paper, will be aimed at extending the proposed Shb-based method to other applicative contexts (e.g., NDT/NDT, GPR investigations, biomedical imaging, or food quality assessment) involving - for instance - differential formulations of the ISP to embed the a-priori knowledge on a reference/healthy background scenario. The exploitation of qualitative imaging algorithms integrated within the PF Block (e.g., time-reversal imaging or the Multiple Signal Classification Method [1]) will be investigated, as well, to generate initial guesses of D and acquire useful information (e.g., number of scatterers, position, and size) enabling a suitable definition of the modeling used for the successive Shb inversion. Finally, the extension of the methodology to deal with 3D-IS scenarios will be pursued by exploiting properly-defined 3D modeling approaches within the PF Block, still based on the available information on the unknown targets.

APPENDIX I

The LHS strategy is implemented through the following procedure:

- Uniformly divide the admissible range $\hat{\alpha}_k = [\alpha_{\min}^{(s)}, \alpha_{\max}^{(s)}]$ of each k-th ($k = 1, ..., K$) DoF into $S_1$ intervals $\left\{ \alpha_k^{(s)} ; s = 1, ..., S_1 \right\}$ such that

$$\alpha_k = \bigcup_{s=1}^{S_1} \alpha_k^{(s)}$$

- For each k-th ($k = 1, ..., K$) variable, randomly choose one value $\alpha_k^{(s)}$ within each s-th ($s = 1, ..., S_1$) interval, $\alpha_k^{(s)}$, and form the corresponding set $S_k = \left\{ \alpha_k^{(s)} ; s = 1, ..., S_1 \right\}$.

- Until $s = S_1$, form the s-th K-dimensional sample $\xi_s^{(s)} = \left\{ \xi_k^{(s)} ; k = 1, ..., K \right\}$ by letting $\xi_k^{(s)} = R(S_k)$ ($k = 1, ..., K$) where the operator $R(\cdot)$. outputs the value of one randomly-chosen entry of $S_k$, which is then removed from it. Update the index $s \leftarrow (s + 1)$ and repeat.

ACKNOWLEDGEMENTS

A. Massa wishes to thank E. Vico for her never-ending inspiration, support, guidance, and help.

REFERENCES

[1] X. Chen, Computational Methods for Electromagnetic Inverse Scatter- ing. Hoboken, NJ, USA: Wiley, 2018.
[2] A. Abubakar, P. M. van den Berg, and J. Mallorqui, “Imaging of biomedical data using a multiplicative regularized contrast source inversion method,” IEEE Trans. Microw. Theory Techn., vol. 50, no. 7, pp. 1761-1771, Jul. 2002.
[3] P. Mojab and J. LoVetri, “Microwave biomedical imaging using the multiplicative regularized Gauss-Newton inversion,” IEEE Antennas Wireless Propag. Lett., vol. 8, pp. 645-648, Jul. 2009.
[4] Y. Gao and R. Zoughi, “Millimeter wave reflectometry and imaging for noninvasive diagnosis of skin burn injuries,” IEEE Trans. Instrum. Meas., vol. 66, no. 1, pp. 77-84, Jan. 2017.
[5] A. Afsari, A. M. Abbosh, and Y. Rahmat-Samii, “Modified Born iterative method in medical electromagnetic tomography using magnetic field fluctuation contrast source operator,” IEEE Trans. Microw. Theory Techn., vol. 67, no. 1, pp. 454-463, Jan. 2019.
[6] X. Song, M. Li, F. Yang, S. Xu, and A. Abubakar, “Study on joint inversion algorithm of acoustic and electromagnetic data in biomedical imaging,” IEEE J. Multiscale Multiphys. Comput. Techn., vol. 4, pp. 2-11, 2019.
[7] T. Cui, W. C. Chew, A. A. Aydiner, and S. Chen “Inverse scattering of two-dimensional dielectric objects buried in a lossy earth using the distorted Born iterative method,” IEEE Trans. Geosci. Remote Sens., vol. 59, no. 2, pp. 339-346, Feb. 2001.
[8] M. Salucci, G. Oliveri, and A. Massa, “GPR prospecting through an inverse scattering-frequency-hopping multi-focusing approach,” IEEE Trans. Geosci. Remote Sens., vol. 53, no. 12, pp. 6573-6592, Dec. 2015.
[9] M. Salucci, L. Poli, and A. Massa, “Advanced multi-frequency GPR data processing for non-linear deterministic imaging,” Signal Proc., vol. 132, pp. 306-318, Mar. 2017.
[10] M. Salucci, L. Poli, N. Anselmi and A. Massa, “Multifrequency particle swarm optimization for enhanced multiresolution GPR microwave imaging,” IEEE Trans. Geosci. Remote Sens., vol. 55, no. 3, pp. 1305-1317, Mar. 2017.
[11] Z. Liu, C. Li, D. Lesselier, and Y. Zhong, “Fast full-wave analysis of damaged periodic fiber-reinforced laminates,” IEEE Trans. Antennas Propag., vol. 66, no. 7, pp. 3540-3547, Jul. 2018.
[12] Z. Zoughi, Microwave Nondestructive Testing and Evaluation. Amsterdam, The Netherlands: Kluwer, 2000.
[13] A. Afsari, A. M. Abbosh, and Y. Rahmat-Samii, “Modified Born iterative method for solving electromagnetic inverse scattering problems of two-dimensional dielectric objects buried in a lossy earth using the distorted Born iterative method,” IEEE Trans. Antennas Propag., vol. 49, no. 12, pp. 1812-1820, Dec. 2001.
[14] K. Xu, Y. Zhong, X. Chen, and D. Lesselier, “A fast integral equation-based method for solving electromagnetic inverse scattering problems with inhomogeneous background,” IEEE Trans. Antennas Propag., vol. 66, no. 8, pp. 4228-4239, Aug. 2018.
[15] Y. Chu, K. Xu, Y. Zhong, X. Ye, T. Zhou, X. Chen, and G. Wang, “Fast microwave through wall imaging method with inhomogeneous background based on Levenberg-Marquardt algorithm,” IEEE Trans. Microw. Theory Techn., vol. 67, no. 3, pp. 1138-1147, Mar. 2019.
