

Using Covariance Matrix Adaptation Evolutionary Strategy to boost the search accuracy in hierarchic memetic computations

Jakub Sawicki^{a,1,*}, Marcin Łoś^{a,1}, Maciej Smółka^{a,1}, Julen Alvarez-Aramberri^{b,c,2}

^a*AGH University of Science and Technology, Kraków, Poland*

^b*Columbia University, Lamont-Doherty Earth Observatory, Palisades, New York, USA*

^c*Basque Centre for Applied Mathematics (BCAM), Bilbao, Spain*

Abstract

Many global optimization problems arising naturally in science and engineering exhibit some form of intrinsic ill-posedness, such as multimodality and insensitivity. Severe ill-posedness precludes the use of standard regularization techniques and necessitates more specialized approaches, usually comprised of two separate stages - global phase, that determines the problem's modality and provides rough approximations of the solutions, and a local phase, which refines these approximations.

In this work, we attempt to improve one of the most efficient currently known approaches - Hierarchic Memetic Strategy (HMS) - by incorporating the Covariance Matrix Adaptation Evolutionary Strategy (CMA-ES) into its local phase. CMA-ES is a stochastic optimization algorithm that in some sense mimics the behavior of population-based evolutionary algorithms without explicitly evolving the population. This way, it avoids, to an extent, the associated cost of multiple evaluations of the objective function.

We compare the performance of the HMS on relatively simple multi-

*Corresponding author

Email addresses: jsawicki@agh.edu.pl (Jakub Sawicki), marcin.los.91@gmail.com (Marcin Łoś), smolka@agh.edu.pl (Maciej Smółka), julen.alvarez.aramberri@gmail.com (Julen Alvarez-Aramberri)

¹The work presented in this paper has been partially supported by Polish National Science Center grants No. DEC-2015/17/B/ST6/01867 and by the AGH statutory research grant.

²This research is supported by the Basque Government through the BERC2018-2021 program and by the Spanish Ministry of Science, Innovation and Universities: BCAM Severo Ochoa accreditation SEV-2017-0718.

modal benchmark problems and on an engineering problem. To do so, we consider two configurations: the CMA-ES and the standard SEA (Simple Evolutionary Algorithm). The results demonstrate that the HMS with CMA-ES in the local phase requires less objective function evaluations to provide the same accuracy, making this approach more efficient than the standard SEA.

Keywords: inverse problems, multi-objective optimization methods, memetic algorithms, ill-posed problems

1. Introduction

Many global optimization problems provided by environmental science, geophysics, medical diagnosis, economy, and technology are not correctly posed. For instance, the ambiguity appearing by lens design is described in [1], while [2] mentions the ill-conditioning of the heat flux estimation in irradiative dryer/furnace. Some examples of uncertain geophysical problems are [3] (calibration of conceptual Rainfall-Runoff Models) and [4] (investigation of oil and gas resources). The problems of cancer diagnosis (see, e.g. [5]) as well as restoring the high-resolution image from the low-level Electrical Impedance Tomography (EIT) images [6] might also be classified as ill-posed.

Generally, the ill-posedness mentioned above is not caused by neglecting the necessary mathematical preciseness but generally, by the insufficient knowledge about the problem's nature (e.g., lack of the total observability, incomplete and inaccurate measurements, computational errors or even poor formal model of the phenomenon under consideration only available).

Typical symptoms of such ill-conditioning are the objective multimodality or/and the regions of objective insensitivity, which might be called "lowlands" in the optimization landscape (see [7] for details). The traditional formulations and solving methods are difficult to apply for such problems. In particular, strategies leading only to find the global, stable minimizer and approximate it by a gradient based method starting from an arbitrary admissible point, are unable to deliver satisfactory information of solutions to such problems.

More reliable knowledge about the problem should contain the information about all local minimizers and the objective behavior in some neighborhood of each of them.

If the ill-conditioning of the problem is mild and has a form of small perturbation (e.g., the problem is conditionally well-posed in the sense of

Tikhonov [8]), then the solution can be obtained by applying a local convex minimization method and some regularized objective supplement (see e.g. [9]).

Unfortunately, global optimization problems with a multimodal and an insensitive objective can rarely be solved by a single convex optimization algorithm. Typical strategies are composed of two phases: the *global phase* in which the objective modality is determined and the *local phase*, in which the solution's approximation is refined in the basin of attraction of each local minimizer (see, e.g., [10, 11]).

The global phase can be performed by a stochastic search, sometimes of a complex structure, involving post processing of the random sample with an artificial intelligence method (see e.g. [12, 13]). It is also possible to apply an evolutionary or memetic approach in the global phase (see [14, 15]). The local phase is traditionally performed by gradient-based methods (see, e.g., [16]), if the total strategy leads to find only one minimizer in each distinguished basin of attraction.

In this paper, we use the Hierarchic Memetic Strategy (HMS), which is a concurrent multi-phase global optimizer [17]. This is a complex strategy aimed at solving ill-posed inverse problems formulated as multi-modal global optimization tasks with expensive objective function evaluations. The goal of HMS is to provide possibly all solutions with satisfactory accuracy and in a reasonable time. The strategy was already tested on several real-world engineering inverse problems and proved its superiority over commonly-used global optimizers like the Simple Evolutionary Algorithm (SEA) and the multi-start method, cf. [18, 4, 17].

Roughly, HMS consists of running in parallel the tree of demes searching with growing accuracy and locality when passing from the root to leaves. The information about the promising search regions propagates down the tree to introduce new child-demes in the best positions found to date. The root and upper demes are responsible for the global phase, while the leaves are involved in the local HMS phase. It is possible to execute both phases concurrently.

While we have applied the Evolutionary Algorithm (EA) with a normal and α -stable mutations [19] in the global HMS phase, we have already checked the following algorithms in the local phase of HMS:

- EA and steepest descent methods based on the accurate and approximated objective gradient (see e.g. [4]),
- the steepest descent methods associated with the local Tikhonov regularization [20],

- EA equipped with the multi-winner selection [21] and the local objective approximation [22], if the insensitivity regions have to be determined.

In this paper, we concentrate on improving the local phase of the hierarchical memetic search by introducing a Covariance Matrix Adaptation Evolutionary Strategy (CMA-ES) as the search engine in the leaves, similarly as in [14]. With this strategy, we aim to:

- Increase the efficiency with respect to the EA strategy by reducing the total number of the most expensive fitness calls in the leaves.
- Reduce the number of parameters for tuning (since CMA-ES includes a self-tuning mechanism).
- Obtain additional information about the local fitness behavior drawn from the parameters (e.g., the mean and the correlation matrix) of the CMA-ES sampling measure, especially in the case, when the basins of attraction of local minimizers significantly vary in shape and size.

The structure of the paper is as follows. Section 2.1 shortly describes the HMS strategy. In Section 2.3 we show how to apply CMA-ES as a deme evolutionary engine in HMS. Section 3 shows the results of the evaluation of HMS-CMA-ES composition on multi-modal benchmark functions and on an engineering problem. We provide a comparison of the new strategy with both the single-population CMA-ES and the HMS with SEA as the deme evolutionary engine. We finish the paper with some conclusions.

2. Solving strategy

2.1. Hierarchic Memetic Strategy

In this section, we describe the algorithmic framework of our global search strategy. In particular, we focus on multi-population evolutionary algorithms.

The HMS is a mixture of a multi-deme evolutionary algorithm and some deterministic techniques, such as gradient-based optimization methods. The HMS algorithm is employed for the accuracy boosting, dynamic accuracy adjustment for the computation time reduction, sample clustering, and additional evolutionary components. Additionally, it is equipped with a multi-winner selection operator for the objective plateau discovery, cf. [17, 21, 22].

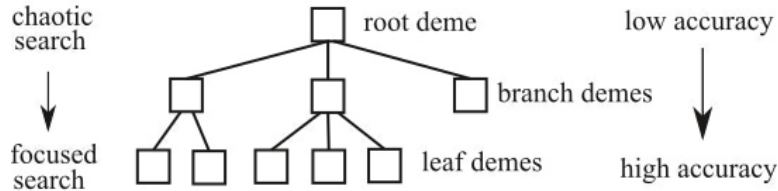


Figure 1: HMS deme structure

In this paper, we focus on the evolutionary core of HMS that was its straightforward predecessor and has been developed under its own name, i.e., Hierarchic Genetic Strategy (HGS), cf. [23, 24, 18, 4].

As a multi-population evolutionary strategy, the HMS makes use of a particular data structure that provides an organization among the sub-populations (demes). It is a tree with a fixed maximal height and variable internal node degree (see Figure 1). The tree has the parent-child semantics. An outline of the whole strategy is presented in Algorithms 1 and 2.

Algorithm 1 The evolutionary core of HMS

Require: Maximal height of the tree $M > 0$

```

1:  $root \leftarrow \text{CREATEDEME}(0, \emptyset)$ 
2:  $ActiveDemes \leftarrow \{root\}$ 
3: repeat
4:   for all  $d \in ActiveDemes$  do
5:     Run metaepoch in  $d$ 
6:     if  $d$  satisfies local stopping condition then
7:        $ActiveDemes \leftarrow ActiveDemes \setminus \{d\}$ 
8:     else if  $level(d) < M$  and  $d$  satisfies sprout condition then
9:        $cb \leftarrow currentBest(d)$ 
10:       $child \leftarrow \text{CREATEDEME}(level(d) + 1, \{cb\})$ 
11:       $d.children \leftarrow d.children \cup \{child\}$ 
12:       $ActiveDemes \leftarrow ActiveDemes \cup \{child\}$ 
13:    end if
14:  end for
15: until global stopping condition is satisfied

```

Each sub-population evolves according to a single-population evolutionary algorithm like the SEA. We use the same evolutionary engine for all demes

Algorithm 2 Function CREATEDEME

```
1: function CREATEDEME(level, initialPoints)
2:    $pop \leftarrow \emptyset$ 
3:   if initialPoints =  $\emptyset$  then
4:     Sample  $pop$  uniformly over the search domain
5:   else if initialPoints =  $\{b\}$  then
6:     Sample  $pop$  with distribution  $\mathcal{N}(b, \sigma(\text{level}))$ 
7:   end if
8:    $d \leftarrow \text{new}$ 
9:    $d.\text{population} \leftarrow pop$ 
10:   $d.\text{level} \leftarrow \text{level}$ 
11:   $d.\text{searchAccuracy} \leftarrow \text{accuracy}(\text{level})$ 
12:  return  $d$ 
13: end function
```

on a particular level of the tree. In a typical setup, the HMS begins with a single parentless deme, i.e., the root. A single HMS global step (called a *metaepoch*) consists in executing a prescribed number of genetic epochs of each active deme. During a metaepoch, the tree structure remains unchanged. However, after that, each active deme located at a non-maximal tree level tries to create a new (next-level) deme by sampling some points around its current best individual. This operation is called *sprouting* and it can change the tree structure by adding new child nodes. Nevertheless, it is conditional: a node cannot sprout a new child if its current best is too close to other demes located at the target tree level. A node is active as long as its local stopping condition is not satisfied. The whole tree evolves until a global stopping condition starts to hold. The range and accuracy of the performed evolutionary search are different among the HMS demes. The search is more chaotic and inaccurate in the root population, and it becomes more and more focused and accurate while going towards the leaves. The reason is because the task of higher-level populations is to discover promising search areas that are subsequently explored more thoroughly by lower-level populations. The final solutions to the whole strategy are then discovered by the leaf demes. For the full description of the HGS and the HMS we refer the reader to papers [23, 24, 17, 21, 25, 22].

2.2. Covariance Matrix Adaptation Evolutionary Strategy

CMA-ES is a stochastic, population-based method of optimization, well suited to the semi-local search in continuous domains $\mathcal{D} \subset \mathbb{R}^n$, $n \geq 1$. The

strategy was introduced and initially tested in [26].

The main idea of CMA-ES is to represent the sampling measure as a multivariate normal distribution $\mathcal{N}(m_k, \sigma_k^2 C_k)$ appearing in each iteration k . It is parametrized by a tuple (m_k, σ_k, C_k) ; where m_k is the mean, σ_k the standard deviation scaling factor (step size), and C_k the normalized covariance matrix.

CMA-ES starts its operation by creating an initial population P_0 by λ -times sampling with return using some initial parameters: the step size σ_0 , a distribution's covariance matrix C_0 of ones at the diagonal and zeros outside, and the mean m_0 chosen in the admissible domain. At each iteration, the tuple of parameters is modified based on the current evaluated population P_k to better fit the central part of the basin of attraction of the nearest minimizer (see (1)).

$$\begin{array}{ccc}
 (k, (m_k, \sigma_k, C_k)) & & \\
 \downarrow & \lambda - \text{times sampling with return} & \\
 & \text{according to the distribution} & \\
 & \mathcal{N}(m_k, \sigma_k^2 C_k) & \\
 P_k & & (1) \\
 \downarrow & \text{deterministic adaptation step} & \\
 (k + 1, (m_{k+1}, \sigma_{k+1}, C_{k+1})) & &
 \end{array}$$

We employ two stopping criteria: the algorithm stops when (a) the parameters $(m_{k+1}, \sigma_{k+1}, C_{k+1})$ do not differ significantly from (m_k, σ_k, C_k) , or when (b) the fitness of at least one individual is below the satisfactory level.

The high efficiency of CMA-ES in the local search in irregular landscapes in comparison with the conventional EA comes mainly from the more synthetic and concise representation of the knowledge about the problem gathered during all iterations (contained in the parameters (m_k, σ_k, C_k)). Moreover, its consecutive improvements need less objective evaluations.

The extraordinary capabilities shown by CMA-ES in the local gradient-free search are the result of an economical and flexible representation of a sampling measure density and the deterministic character of the adaptation technique used throughout the strategy. The topic is broadly covered in the literature, in particular in [26, 27, 28, 14].

2.3. Improving HMS by the CMA-ES component

We use CMA-ES only in leaves of the HMS hierarchy. The higher levels of the computation tree perform a more global search for which the sampling

distribution adaptation offered by SEA is better suited and can be controlled by the mutation and crossover parameters. We argue that CMA-ES with its local-like capabilities is going to be more effective than ordinary SEA at converging to local and global minimizers if the starting population is located in their basins of attraction.

3. Testing new component

The benchmarks shown here are meant to test the CMA-ES algorithm coupled with HMS applied to solve multimodal problems.

Benchmark functions. In tests 1-4 we use two benchmark functions defined on 2- and 4-dimensional domains. Both are composed of inverted Gaussian functions. The function is defined as:

$$g_{x^0}^r(x) = 1 - \exp(-(x - x^0)^T S(x - x^0)) \quad , \quad (2)$$

where $x^0 \in \mathcal{D}$ is a point in the domain, $r \in \mathbb{R}_+$ is a constant, and S is a diagonal matrix of size n with diagonal elements equal to r^{-1} and zeros outside, and n is the number of dimensions of the domain \mathcal{D} . The function $g_{x^0}^r$ has a minimum of 0 at point x^0 and raises to 1 exponentially away from x^0 . Function $g_{x^0}^r$ can be multiplied by another function of the same type to obtain multimodal expressions.

For example, the two-dimensional benchmark used in Tests 1-3 is defined as follows:

$$f_1(x) = g_{(1,1)}^1(x) \cdot g_{(6,1)}^{0.2}(x) \cdot g_{(7,8)}^{0.2}(x) \quad x \in [0, 10]^2 \quad , \quad (3)$$

and is presented on Figure 2. It has minima in (1, 1), (6, 1) and (7, 8).

The four-dimension benchmark used in Test 4 is defined by:

$$f_2(x) = g_{(2,2,2,2)}^2(x) \cdot g_{(6,1,2,3)}^{0.4}(x) \cdot g_{(7,8,2,3)}^{0.4}(x) \quad x \in [0, 10]^4 \quad . \quad (4)$$

It has minima in (2, 2, 2, 2), (6, 1, 2, 3) and (7, 8, 2, 3). The radii were scaled up by 2 compared to f_1 , in order to maintain a similar susceptible region.

We describe the engineering problem that we employ in Test 6 in Section 3.6.

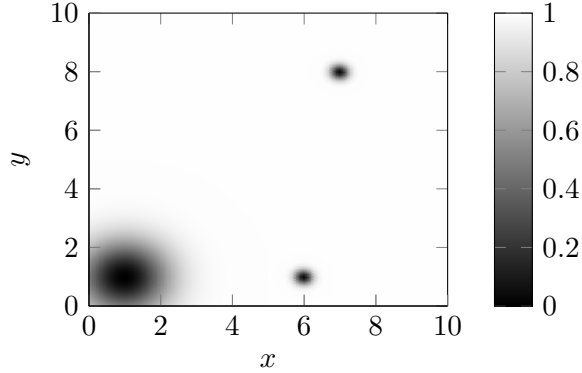


Figure 2: The multimodal 2D benchmark function.

Metrics. We assess the algorithms with the following four metrics: *covered minima count* (CMC), *M covered minima evaluations* (*M-CME*), *mean minimum distance* (MMD) and *mean minimum distance fitness* (MMDF).

The number of minima covered (CMC) is calculated by taking all the points generated by the demes (referenced from now on as deme points) and finding the closest one for each minimum. If the point is within a prescribed distance around the minimum, then we consider the minimum covered.

The *M-CME* metric provides the number of evaluations that an algorithm needs to cover *M* minima (in our case between one and three) using the idea of the *first hitting time*.

M-CME is equal to (*M* - 1)-*CME* plus the number of evaluations required to cover an additional minimum. The value of *M-CME* is undefined for a particular run of an algorithm. Besides, it is only included in the statistics when the run is unable to cover all *M* minima. Nevertheless, the statistics can be evaluated if needed.

MMD is computed as follows for each benchmark run:

$$\text{MMD} = \frac{1}{3} \sum_{i=1}^3 \inf_{x \in P_{\text{leaf}}} \|x - m_i\|, \quad (5)$$

where P_{leaf} is an aggregated set of points from the leaf-demes, m_i is the center of the i^{th} minimum and by $\|\cdot\|$ we mean Euclidean norm. Unlike the CMC, MMD is parameterless.

MMDF is computed similarly to MMD:

$$\text{MMDF} = \frac{1}{3} \sum_{i=1}^3 f \left(\arg \min_{x \in P_{\text{leaf}}} \|x - m_i\| \right) - f(m_i). \quad (6)$$

In this case, we also consider the points from the leaves that are closest to the minima, and we average their fitness.

Configurations. Table Table 1 includes the parameters used for each of the tests. We select these values based on previous works solving similar problems (see, e.g., [17]).

3.1. Test 1 — comparison of HMS-SEA, CMA-ES and HMS-CMA-ES

Here we test the performance of the CMA-ES, the HMS-CMA-ES and the HMS-SEA using the CMC metric and the total number of evaluations. We check the performance of the CMA-ES and we compare it to HMS versions in terms of the number of evaluations. It is the only test which takes the single-deme CMA-ES into account.

We run each configuration 50 times and we show the results in Table 2. CMA-ES alone is unable to find more than one minimum, as expected — it was included to compare the number of evaluations needed. HMS-SEA is inefficient at covering more than one minimum and it uses 9 times larger budget than HMS-CMA-ES. HMS-CMA-ES covers more minima than the HMS-SEA, however, it still is not able to do it with great confidence — it fails to localize any minima in some cases. The budget of HMS-CMA-ES is larger than a single-deme CMA-ES, but the overhead is justified in better performance as far as CMC metric goes.

3.2. Test 2 — M-CME comparison

We wanted to check how would using CMA-ES in the lowest level of the HMS hierarchy modify the performance of such strategy. We aim at studying the impact of using the CMA-ES strategy in the lowest level of the HMS.

We run each variant 100 times to collect statistics of the 1-, 2- and 3-CME metrics and the number of runs which succeeded in covering that many minima.

The results are presented in Table 3. The performance of both variants is comparable though. HMS-CMA-ES does better when min and max statistics are taken into account, but the median in the 3-CME is larger than for HMS-SEA variant.

3.3. Test 3 — MMD and MMDF in function of budget

To analyze the performance of the strategy depending on the budget, we set up two benchmarks and we run them with different budgets. They have global stopping conditions of their own, so they may actually have used

		ME	Level 1	Level 2
Test 1	HMS-SEA	3	SEA/20/0.1/0.5/5 NS/5 SC/0.9/0.8	SEA/20/0.7/0.5/0.4 NC/10 ⁻³ /5
	CMA-ES	3	CMA-ES/10/1	—
	HMS-CMA-ES	3	SEA/30/0.1/0.5/5 NS/5 SC/0.9/2	CMA-ES/10/1 NC/10 ⁻⁶ /5
Test 2	HMS-SEA	3	SEA/20/0.1/0.5/5 NS/10 SC/0.9/2	SEA/10/0.7/0.5/0.4 NC/10 ⁻³ /5
	HMS-CMA-ES	3	SEA/20/0.1/0.5/5 NS/10 SC/0.9/2	CMA-ES/10/1 NC/10 ⁻⁶ /5
Test 3	HMS-SEA	3	SEA/20/0.1/0.5/5 NS/10 SC/0.9/1	SEA/10/0.7/0.5/0.4 NC/10 ⁻³ /5
	HMS-CMA-ES	3	SEA/20/0.1/0.5/5 NS/10 SC/0.9/1	CMA-ES/10/1 NC/10 ⁻⁶ /5
Test 4	HMS-SEA	3	SEA/100/0.1/0.5/5 NS/10 SC/0.9/1	SEA/10/0.7/0.5/0.4 NC/10 ⁻³ /5
	HMS-CMA-ES	3	SEA/100/0.1/0.5/5 NS/10 SC/0.9/1	CMA-ES/10/1 NC/10 ⁻⁶ /5
Test 5	HMS-SEA	3	SEA/100/0.1/0.5/5 NS/10 SC/0.9/1	SEA/10/0.7/0.5/0.4 NC/10 ⁻³ /5
	HMS-CMA-ES	3	SEA/100/0.1/0.5/5 NS/10 SC/0.9/1	CMA-ES/10/1 NC/10 ⁻³ /5
Test 6	HMS-SEA	1	SEA/10/0.1/0.5/1 SC/3.5/0.05	SEA/8/0.2/0.4/0.05 NC/0.1/3
	HMS-CMA-ES	1	SEA/10/0.1/0.5/1 SC/3.5/0.05	CMA-ES/8/0.1 NC/0.1/3

Table 1: Configurations for algorithms used in the tests. ME stands for metaepoch length. Each level is described with three components. First, algorithm type: SEA or CMA-ES. SEA parameters are: population size, crossover probability, mutation probability, and mutation standard deviation. CMA-ES parameters are: population size and initial step size. Second, local stop condition: no sprout (NS) or no change (NC). NS parameter is the number of metaepochs. NC parameters are: minimal change and the number of epochs. Third, sprouter configuration (SC): fitness threshold and allowed distance to other demes and sprouts.

Algorithm	CMC			Evaluations		
	mean	min	max	mean	min	max
HMS-SEA	1.0 ± 0.7	0	2	9664 ± 1044	5015	10 202
CMA-ES	1.0	1	1	669 ± 45	591	781
HMS-CMA-ES	1.8 ± 0.8	1	3	1043 ± 384	580	2139

Table 2: Results of Test 1. Each row corresponds to the algorithm’s variant tested. The mean CMC and the number of evaluations taken by each variant are presented, along with the standard deviation of the mean, and minimal and maximal values attained.

Algorithm	1-CME		2-CME		3-CME	
	count	min/med/max	count	min/med/max	count	min/med/max
HMS-SEA	100	90/150/1350	81	100/540/1710	40	370/960/2640
HMS-CMA-ES	100	81/141/423	83	82/382/1053	42	253/1123/1683

Table 3: Results of Test 2. Each row corresponds to the algorithm’s variant tested. 1-CME, 2-CME and 3-CME metrics are presented (minimum, median and maximum value). Moreover, for each metric a number of runs which managed to cover that number of minima (per 100) is shown.

smaller number of evaluations. We used CMC, MMD and MMDF metrics to compare the variants.

We consider a range of budgets from 200 to 5000 and for each budget we run 100 repetitions of each algorithm.

Figure 3 shows a significant improvement for budgets larger than 1000 evaluations when the MMD and MMDF metrics are employed. The number of minima covered is slightly better for the CMA-ES.

The sub-optimal peak performance is not a problem with CMA-ES specifically. While CMA-ES is expected to improve the efficiency on the leaf-level, the global search capabilities of the strategy are mostly on the side of HMS. Further configuration of HMS may allow for better CMC results, but it is beyond the scope of this test.

The CMC metric for both variants is around 2.4. However, HMS-SEA required 4×10^3 evaluations to reach it while HMS-CMA-ES took only 1.5×10^3 — over 2.5 times less.

3.4. Test 4 — MMD and MMDF in function of budget for 4D benchmark

To study the strategy on a more demanding problem, we run this test on a 4D benchmark (for the function f_2 defined in (4)). Most of the parameters of the algorithms remain unaltered. Since the 4D domain is more difficult,

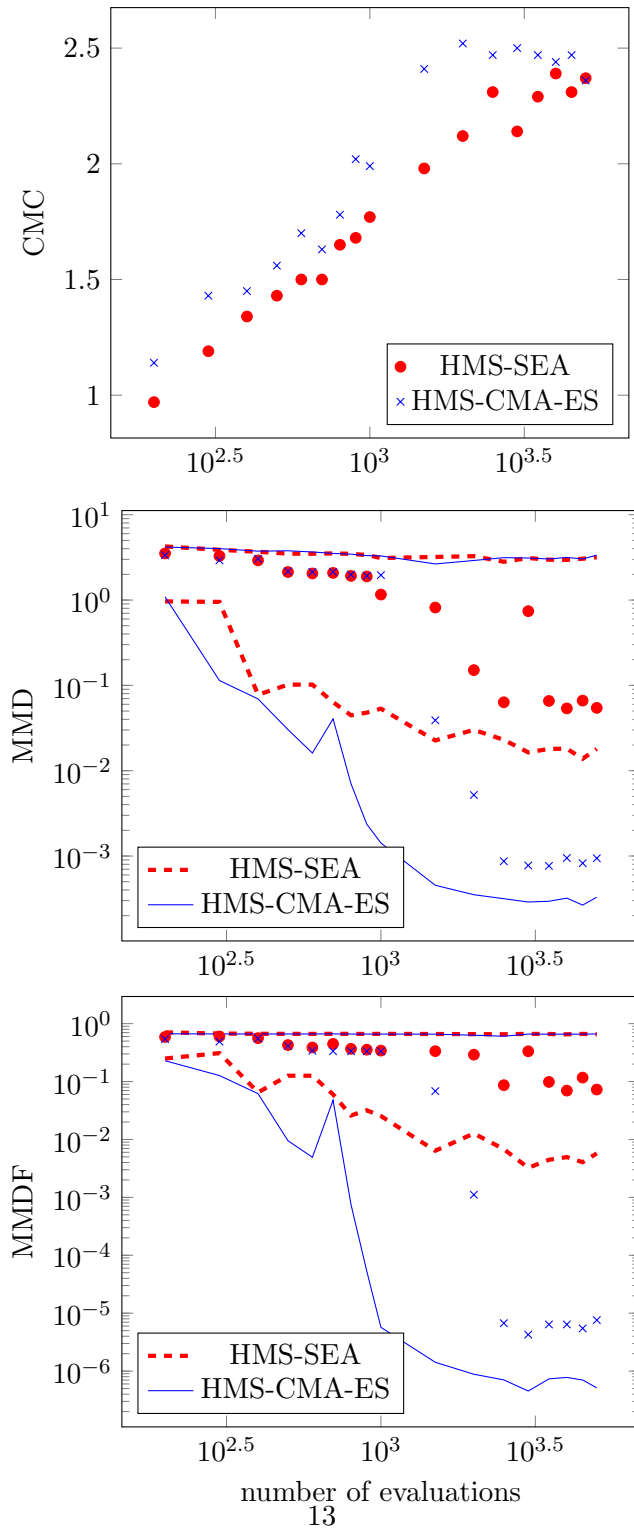


Figure 3: Results of Test 3. All the values shown by points are medians of the results. For the two last metrics, 5. and 95. percentiles are shown as lines as well.

Algorithm	1-CME		2-CME		3-CME	
	count	min/med/max	count	min/med/max	count	min/med/max
HMS-SEA	100	430/6610/62240	54	3090/77915/306320	11	74010/168710/327650
HMS-CMA-ES	100	406/1092/2511	100	642/30038/144483	87	16008/71427/155035

Table 4: Results of Test 5. Each row corresponds to the algorithm’s variant tested. 1-CME, 2-CME and 3-CME metrics are presented (minimum, median and maximum value). Moreover, for each metric a number of runs which managed to cover that number of minima (per 100) is shown.

we set the size of the population to 100 to enhance the exploration. The budget ranges from 10^3 to 10^6 evaluations. We run 100 repetitions of each algorithm variant per budget.

The results, which include the same metrics as in the previous test, are shown in Figure 4. The first panel of this figure shows that HMS-CMA-ES does better than its counterpart, confidently beginning to cover all the minima around 10^5 evaluations. When the number of budgets is sufficiently large (10^4 or more), the gains of HMS-CMA-ES increase.

The CMC metric value for HMS-SEA stabilizes around 1.8 and takes at least 10^5 evaluations. HMS-CMA-ES achieved this accuracy in 10^4 , which is 6 times less than the SEA variant.

3.5. Test 5 — *M-CME comparison for 4D benchmark*

We run each variant 100 times to collect statistics of the 1-, 2- and 3-CME metrics and the number of runs which succeeded in covering that many minima. The setup comparison is the same as in Section 3.2.

We present the results in Table 4. In this case, HMS-CMA-ES variant is able to identify the minima for a significantly smaller budget than HMS-SEA.

3.6. Test 6 — *Magnetotelluric (MT) oil exploration method*

The MT method [29] is a geophysical prospecting method governed by Maxwell’s equations and intended to determine the conductivity (or resistivity) distribution of the Earth’s subsurface. This method employs natural sources, while the receivers are placed at the surface of the Earth or the seafloor. Being clean and inexpensive, it has been applied for different purposes, such as the mapping of active faults [30], the study of volcanoes [31], or the exploration of offshore hydrocarbons [32].

Figure 5 corresponds to the considered Earth formation: a layered model with an embedded inhomogeneity. The computational domain then consists

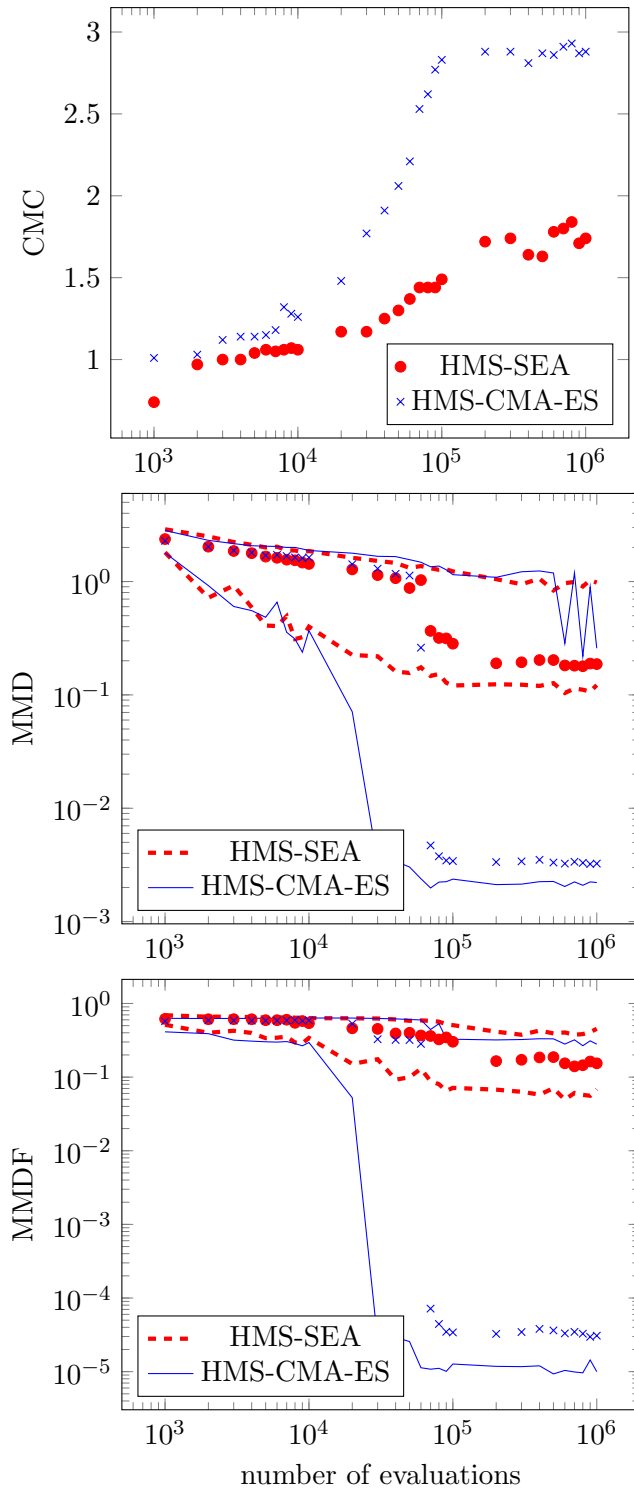


Figure 4: Results of Test 4. All the values shown by points are medians of the results. For the two last metrics, 5. and 95. percentiles are shown as lines as well.

of this subsurface formation, the air, a natural source placed at the ionosphere, and the receivers located at the surface of the Earth. We truncate the computational domain with the Perfectly Matched Layer (PML) described in [33] and by setting Dirichlet homogeneous boundary conditions.

To solve the equations involved in this problem, in the present work we employ an adaptive Finite Element Method described in [34].

A thorough introduction of the problem has been performed by Smořka et al. [17, Sec. 3]. The task is to identify the conductivity $\sigma = \{\sigma_i\}_{i=1,\dots,4}$ of the regions presented in Figure 5. The domain of the problem $\mathcal{D} = [0, 6]^4$ is mapped to the actual conductivity by a function $\mathcal{D} \ni x = \{x_i\}_{i=1,\dots,4} \rightarrow \{10^{x_i-1}\}_{i=1,\dots,4} = \sigma$.

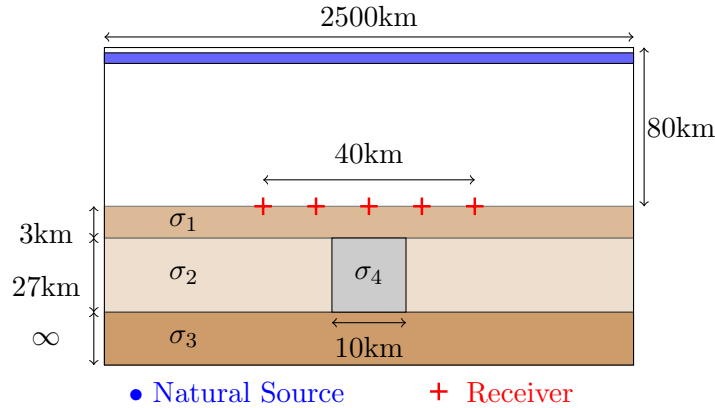


Figure 5: The domain of MT problem

The fitness values are transformed by a function

$$\mathbb{R}_+ \ni y \rightarrow \log_{10}(y/y_0) \in \mathbb{R}_+ \cup \{0\}, \quad (7)$$

where y_0 is the original fitness value returned by the MT solver at the point $(1, 2, 10, 3)$ in the conductivity space.

We compare HMS-SEA and HMS-CMA-ES algorithm variants. Their performance is evaluated by measuring the best fitness that their leaf demes are able to identify. We run them 10 times with a fixed budget of 1000 evaluations.

The results are presented in Figure 6. Using the same number of evaluations, HMS-CMA-ES is able to approach the global optimum fitness value closer.

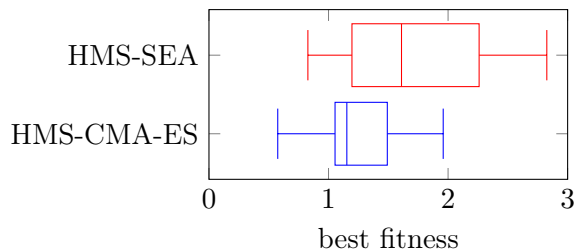


Figure 6: Results of Test 6. For each variant minimum, 1st quartile, median, 3rd quartile and maximum are shown for the best fitness metric calculated from 10 runs.

4. Conclusions

The paper presents a global search strategy (HMS), proven to be efficient when dealing with ill-posed inverse parametric problems, together with the CMA-ES algorithm.

The HMS operates at several levels of search accuracy. It is known that when the search in the leaves is efficient, the strategy can explore larger regions of the domain (in the search for solutions) at the same computational cost level. Therefore, the objective of this work is to improve the efficiency of the optimization at the finest-search level, that is, in the leaves of the HMS tree.

The HMS strategy (equipped with local gradient-based optimizers) was already tested on real-world engineering problems, cf. [17], and was shown capable of finding almost twice as many good-quality solutions than the multi-start method. At the same time, it was better at avoiding poor-quality solution basins of attraction. However, the SEA with the same budget was unable to recover good-quality solutions.

This paper shows with different numerical examples that the performance of the proposed strategy (the core of the HMS strategy with the CMA-ES in the leaves) is better than when using the SEA in the leaves. In particular, the search is more efficient, in some examples reducing the number of evaluations required to achieve similar performance 9 times.

During its run, the proposed strategy acquires information about the objective function, which is expressed in the CMA-ES sampling distribution. Therefore, an additional benefit that we will explore in the future is to use this information to analyze how the fitness function behaves around its optima. This way, we will be able to determine regions of insensitivity around the solutions of the inverse problem, obtaining, accordingly, more information about the ill-posedness.

Bibliography

References

- [1] M. Isshiki, D. Sinclair, S. Kaneko, Lens design: Global optimization of both performance and tolerance sensitivity, in: *International Optical Design*, Optical Society of America, 2006, p. TuA5. doi:10.1364/IODC.2006.TuA5. URL <http://www.osapublishing.org/abstract.cfm?URI=IODC-2006-TuA5>
- [2] A. Mirsepahi, A. Mehdizadeh, L. Chen, B. O'Neill, M. Mohammadzakeri, Comparison of inverse modelling and optimization-based methods in the heat flux estimation problem of an irradiative dryer/furnace, *Journal of Computational Science* 19 (Supplement C) (2017) 77 – 85. doi:10.1016/j.jocs.2017.01.007.
- [3] Q. Duan, S. Sorooshian, V. Gupta, Effective and efficient global optimization for conceptual rainfall-runoff models, *Water Resource Research* 28 (4) (1992) 1015–1031.
- [4] M. Smolka, E. Gajda-Zagórska, R. Schaefer, M. Paszyński, D. Pardo, A hybrid method for inversion of 3D AC logging measurements, *Applied Soft Computing* 36 (2015) 422–456.
- [5] M. Paruch, E. Majchrzak, Identification of tumor region parameters using evolutionary algorithm and multiple reciprocity boundary element method, *Engineering Applications of Artificial Intelligence* 20 (5) (2007) 647–655.
- [6] A. Abbasi, B. V. Vahdat, A non-iterative linear inverse solution for the block approach in eit, *Journal of Computational Science* 1 (4) (2010) 190 – 196. doi:10.1016/j.jocs.2010.09.001.
- [7] M. Łoś, M. Smolka, R. Schaefer, J. Sawicki, Misfit landforms imposed by ill-conditioned inverse parametric problems, *Computer Science* 19 (2018). doi:10.7494/csci.2018.19.2.2781.
- [8] A. N. Tikhonov, A. Goncharkii, Stepanov, V.V., A. G. Yagola, *Numerical Methods for the Solution of Ill-Posed Problems*, Springer-Verlag, 1995.
- [9] L. Beilina, M. V. Klibanov, *Approximate Global Convergence and Adaptivity for Coefficient Inverse Problems*, Springer, 2012. doi:10.1007/978-1-4419-7805-9.

- [10] R. Horst, P. Pardalos, Handbook of Global Optimization., Kluwer, 1995.
- [11] G. R. Wood, Z. B. Zabinsky, Stochastic adaptive search, in: Handbook of Global Optimization Vol. 2, Kluwer, 2002.
- [12] C. G. E. Boender, A. H. G. Rinnooy-Kan, L. Stougie, G. T. Timmer, A stochastic method for global optimization, Mathematical Programming 22 (1982) 125–140.
- [13] A. H. G. Rinnooy-Kan, G. T. Timmer, Stochastic global optimization methods. part 1: Clustering methods., Mathematical Programming 39 (1987) 27–56.
- [14] M. Preuss, Multimodal Optimization by Means of Evolutionary Algorithms, Natural Computing, Springer, 2015.
- [15] R. Schaefer, Foundation of Genetic Global Optimization, with Chapter 6 by Telega H., Studies in Computational Intelligence Series 74, Springer, 2007.
- [16] A. Törn, A search clustering approach to global optimization, in: L. C. W. Dixon, G. P. Szegö (Eds.), Towards Global Optimisation 2, North-Holland, Amsterdam, 1978, pp. 49–62.
- [17] M. Smółka, R. Schaefer, M. Paszyński, D. Pardo, J. Álvarez-Aramberri, An agent-oriented hierarchic strategy for solving inverse problems, International Journal of Applied Mathematics and Computer Science 25 (3) (2015) 483–498. doi:10.1515/amcs-2015-0036.
- [18] E. Gajda-Zagórska, R. Schaefer, M. Smółka, M. Paszyński, D. Pardo, A hybrid method for inversion of 3D DC logging measurements, Natural Computing 14 (3) (2014) 355–374. doi:10.1007/s11047-014-9440-y. URL <http://dx.doi.org/10.1007/s11047-014-9440-y>
- [19] A. Obuchowicz, M. Smółka, Application of α -stable mutation in a hierarchic evolutionary inverse solver, Journal of Computational Science 17 (2016) 261–269. doi:10.1016/j.jocs.2015.12.006.
- [20] M. Smółka, R. Schaefer, D. Pardo, J. Álvarez-Aramberri, Local Tikhonov regularization in the hierarchic memetic inverse solver, in: Proc. of ECCOMAS Int. Conf. IPM 2017 on Inverse Problems in Mechanics of Structure and Materials, Rzeszów University of Technology Press, 2017, pp. 55–56.

- [21] P. Faliszewski, J. Sawicki, R. Schaefer, M. Smółka, Multiwinner voting in genetic algorithms, *IEEE Intelligent Systems* 32/1 (2017) 40–48. doi:10.1109/MIS.2017.5.
- [22] J. Sawicki, M. Łoś, M. Smółka, R. Schaefer, J. Álvarez-Aramberri, Approximating landscape insensitivity regions in solving ill-conditioned inverse problems, *Memetic Computing* 10 (2018) 279–289. doi:10.1007/s12293-018-0258-5.
- [23] R. Schaefer, J. Kołodziej, Genetic search reinforced by the population hierarchy, in: *Foundations of Genetic Algorithms 7*, Morgan Kaufman, 2003, pp. 383–399.
- [24] B. Wierzba, A. Semczuk, J. Kołodziej, R. Schaefer, Hierarchical Genetic Strategy with real number encoding, in: *Proceedings of the 6th Conference on Evolutionary Algorithms and Global Optimization*, 2003, pp. 231–237.
- [25] E. Gajda-Zagórska, R. Schaefer, M. Smółka, D. Pardo, J. Álvarez-Aramberri, Multi-objective memetic inverse solver reinforced by local optimization methods., *Journal of Computational Science* 18 (2017) 85–94. doi:10.1016/j.jocs.2016.06.007.
- [26] N. Hansen, A. Ostermeier, Completely derandomized self-adaptation in evolution strategies, *Evolutionary computation* 9 (2) (2001) 159–195.
- [27] O. Shir, T. Bäck, Niche radius adaptation in the cma-es niching algorithm, in: *Parallel Problem Solving from Nature - PPSN IX, 9th International Conference, Lecture Notes in Computer Science, Vol. 4193*, Springer, 2006, pp. 142–151.
- [28] O. Shir, E. Emmerich, T. Bäck, Adaptive niche radii and niche shapes approaches for niching with the cma-es, *Evolutionary computation* 18 (1) (2010) 97–126.
- [29] A. Chave, A. Jones, *The magnetotelluric method: Theory and practice*, Cambridge University Press, 2012.
- [30] W. Lestari, A. Widodo, D. Warnana, F. Syaifuddin, W. Utama, J. Rochman, Mapping of kendeng thrust active fault in east java using magnetotelluric method, in: *EAGE-HAGI 1st Asia Pacific Meeting on Near Surface Geoscience and Engineering*, 2018.

- [31] G. Hill, P. Wannamaker, J. Stodt, M. Unsworth, V. Maris, P. Bedrosian, E. Wallin, M. Kordy, Y. Ogawa, P. Kyle, et al., Imaging the magmatic system of erebus volcano, antarctica using the magnetotelluric method, in: AGU Fall Meeting Abstracts, 2017.
- [32] M. Zhdanov, L. Wan, A. Gribenko, M. Čuma, K. Key, S. Constable, Large-scale 3d inversion of marine magnetotelluric data: Case study from the gemini prospect, gulf of mexico, *Geophysics* 76 (1) (2011) F77–F87.
- [33] J. Álvarez-Aramberri, D. Pardo, H. Barucq, A secondary field based hp-finite element method for the simulation of magnetotelluric measurements, *Journal of Computational Science* 11 (2015) 137–144.
- [34] J. Álvarez-Aramberri, D. Pardo, Dimensionally adaptive hp-finite element simulation and inversion of 2d magnetotelluric measurements, *Journal of Computational Science* 18 (2017) 95–105.