

# Asymptotic Guarantee of Success of the *hp*-HGS Strategy

Robert Schaefer<sup>1</sup>, Barbara Barabasz<sup>2</sup> and Maciej Paszyński<sup>1</sup>

<sup>1</sup> Department of Computer Science,

<sup>2</sup> Department of Applied Computer Science and Modeling,

University of Science and Technology, Cracow, Poland,

email: schaefer@agh.edu.pl, paszynsk@agh.edu.pl, barabasz@metal.agh.edu.pl

## Abstract.

We presented the new *hp*-HGS (*hp* adaptive FEM, Hierarchical Genetic Strategy) multi-deme, genetic strategy which can be used for solving parametric inverse problems formulated as the global optimization ones. Its efficiency follows from the coupled adaptation of accuracy derived from the proper balance between the accuracy of *hp*-FEM used for solving direct problem and the accuracy of solving optimization problem. It is shown, that *hp*-HGS can find at least the same set of local extremes as the Simple Genetic Algorithm (SGA). Moreover, the results of asymptotic analysis that verify much less computational cost of *hp*-HGS are recalled from the previous papers.

## 1 Introduction

The parametric inverse problems in which we try to find vector of parameters of the partial differential equation, formulated as the optimal control ones plays a very important role in technology. They are often used in the flaw detection, detection of natural resources and optimal design. Solving such problems is hard, because of their exceptionally high computational cost. It depends on the number of calls of direct problem solver performed by the optimization algorithm, while the computational cost of the direct problem solver depends significantly on the assumed accuracy.

We introduced the new twin adaptive *hp*-HGS strategy for solving such kind of inverse problems (see Schaefer, Barabasz, Paszyński [5]). It uses the *hp*-adaptive Finite Element Method (*hp*-FEM) for solving direct problem, and Hierarchical Genetic Strategy (HGS) for solving the optimal control one. This strategy properly balances the error of solving both problems in order to decrease the total computational cost.

Almost all strategies based on the standard genetic mechanisms that try to make the computation more efficient become risky because of the potential loss of the asymptotic guarantee of success (see e.g. Guus, Boender, Romeijn [3]). The main target of this paper is to presented the formal analysis of *hp*-HGS that proves that this strategy is not worse in finding global and local extremes than the Simple Genetic Algorithm (SGA). Next we recall our results that make possible to compare its expected computational cost to the cost of single population SGA and HGS.

## 2 Problem formulation

We consider the class of parametric inverse problems for which the energy functional can be defined. The direct problem is given by the following variational problem: Find  $u \in u_0 + V$  such that

$$b(d; u, v) = l(v) \quad \forall v \in V \quad (1)$$

where  $u_0$  is the shift of the Dirichlet boundary conditions and  $V$  is the proper Sobolev space. The form of functionals  $b, l$  depend of the modeled physical phenomena (e.g. the variational equations of the linear elasticity described in [1]) and of its parameter  $d \in \mathcal{D}$ , where  $\mathcal{D}$  is the regular compact in  $\mathbb{R}^N$ ,  $N < +\infty$ . For  $b$  symmetric and positively defined, the energy functional takes a form

$$E(d; u) = \frac{1}{2}b(d; u, u) - l(u). \quad (2)$$

The above direct problem can be solved by  $hp$ -adaptive Finite Element Method ( $hp$ -FEM) (see Demkowicz [2]). This method creates the sequence of nested meshes, which are used for solving finite dimensional problems with the increasing accuracy. This meshes are described by the two parameters:  $h$  which is related to the size of elements and  $p$  that stands for the degree of the approximation polynomial. If we have established the computational mesh at the particular step of this method, then the next-step fine mesh is constructed by breaking all elements and increasing the order of polynomial. The problem (1) may be approximated by using  $hp$ -FEM with the finite dimensional subspace  $V_{h,p} \subset V$ , then we intend to find  $u_{h,p} \in u_0 + V_{h,p}$  so that

$$b(d; u_{h,p}, v_{h,p}) = l(v_{h,p}) \quad \forall v_{h,p} \in V_{h,p}. \quad (3)$$

This coarse mesh solution  $u_{h,p}$  is computed in each step of this strategy. The fine mesh solution  $u_{\frac{h}{2}, p+1} \in V_{\frac{h}{2}, p+1}$  that satisfies the equation similar to (3) in the space  $V_{\frac{h}{2}, p+1}$  is computed too. Notice that both approximate solution spaces satisfy  $V_{h,p} \subset V_{\frac{h}{2}, p+1} \subset V$ . Formally, the solutions to (1), (3) depend also on the parameter  $d \in \mathcal{D}$ . The resulted next-step mesh is obtained on the base of the relative  $hp$ -FEM error analysis (see Demkowicz [2]) given by

$$err_{FEM}(d) = \left\| u_{h,p}(d) - u_{\frac{h}{2}, p+1}(d) \right\|_E \quad (4)$$

where  $\|\cdot\|_E$  stands for the energy norm on the space  $V$  (see e.g. Ciarlet [1]). The modifications performed in the fine mesh remains only in elements where the value (4) is large.

We generally assume, that the energy  $J(\hat{d}) = E(\hat{d}; u)$  of the exact solution  $u \in V$  to (1) is known from the experiment and we are looking for the unknown parameter  $\hat{d}$ . Assuming that the proper convergence conditions are satisfied (see Demkowicz [2]) the inverse problem can be formulated as follows

Find  $\hat{g} \in \mathcal{D}$  such that :

$$\lim_{h \rightarrow 0, p \rightarrow +\infty} \left| J_{h,p}(\hat{g}) - J(\hat{d}) \right| \leq \lim_{h \rightarrow 0, p \rightarrow +\infty} \left| J_{h,p}(g) - J(\hat{d}) \right| \quad (5)$$

where  $\hat{g}, g \in \mathcal{D}$  are the approximate parameters and  $J_{h,p}(g) = E(g; u_{h,p}(g))$  is the energy of solution  $u_{h,p}$ .

Algorithm 1: Pseudo-code of the  $j$ -th order deme  $P$  in the  $hp$ -HGS tree.

```

1: if ( $j = 1$ ) then
2:   initialize the root deme;
3: end if
4:  $t \leftarrow 0$ ;
5: repeat
6:   if (global_stop_condition received) then
7:     STOP;
8:   end if
9:   for ( $i \in P^t$ ) do
10:    solve the direct problem for  $g = code(i)$  on the coarse and fine FEM meshes;
11:    compute  $err_{FEM}(g)$  according to the formula (4);
12:    while ( $err_{FEM}(g) > Ratio * \delta_j$ ) do
13:      execute one step of  $hp$  adaptivity;
14:      solve the problem on the new coarse and fine FEM meshes;
15:      compute  $err_{FEM}(g)$  according to the formula (4);
16:    end while;
17:    compute fitness  $f_j(i)$  using the FEM mesh finally established;
18:  end for
19:  if ( $j > 1$ ) then
20:    compute the phenotypes' average and send it to the parental deme;
21:    if (branch_stop_condition( $P^t$ )) then
22:      STOP;
23:    end if
24:  end if
25:  if ( $((t \bmod K) = 0) \wedge (j < m)$ ) then
26:    distinguish the best fitted individual  $x$  from deme  $P^t$ ;
27:    if ( $\neg children\_comparison(x)$ ) then
28:      sprout;
29:    end if
30:  end if
31:  perform proportional selection, obtaining multiset of parents;
32:  perform SGA genetic operations on the multiset of parents;
33:   $t \leftarrow t + 1$ ;
34: until (false)

```

### 3 $hp$ -HGS definition

The idea of  $hp$ -HGS is based on the economic HGS strategy introduced by Kołodziej and Schaefer [4], [7], [6]. Its main idea is running a set of dependent evolutionary demes in parallel. The dependency relation has a tree structure with a restricted number of levels  $m$ . The demes of lower order (close to the root of the structure) perform more chaotic search with the lower accuracy. They only detect the promising regions of the optimization landscape, in which more accurate demes of higher order are activated. The different precision in demes of each level is obtained by the binary genotypes of different length. There is a coherency in search between demes of different order in  $hp$ -HGS tree,

thanks to the special kind of hierarchical, nested encoding that forms the sequence of nested phenotype's grids. The maximum diameter of the phenotype's grid  $\delta_j$ , associated with the demes of the order  $j$  determines the search accuracy at this level in  $hp$ -HGS tree. Of course  $\delta_1 > \dots > \delta_m$ .

Each deme works as SGA and after constant number of genetic epochs  $K$  (called *metaepoch*) each of them, excepting demes of  $m$  order (called *leaf*), sprouts new child-deme in the surrounding of the best fitted individual. The *children\_comparison* function protects to sprout the new deme in the landscape region which is occupied by another deme that has the same parent.

The fitness function  $f_j(i)$  for demes of  $j$ -th order is computed by the  $hp$ -FEM and is based on the energy error  $e_{h,p}(g) = |J_{h,p}(g) - J(\hat{d})|$ , where  $J(\hat{d})$  is the known, real energy and  $g$  represent the parameter decoded from the genotype  $i$  appears in deme of the level  $j$  in  $hp$ -HGS tree. The adaptation of errors is made using the formula (see formula (8) in [10])

$$e_{\frac{h}{2},p+1}(g) \leq \|u_{\frac{h}{2},p+1}(g) - u_{h,p}(g)\|_E^2 + \|u(g) - u_{h,p}(g)\|_E^2 + L |g - \hat{d}| \quad (6)$$

where  $e_{\frac{h}{2},p+1}(g) = |J_{\frac{h}{2},p+1}(g) - J(\hat{d})|$ ,  $L$  stands for the Lipschitz constants of the functionals  $J$  and  $|g - \hat{d}|$  is the error of the inverse problem solution that characterizes the individuals belonging to the HGS demes of the  $j$ -th order. It is easy to observe, that  $|g - \hat{d}|$  corresponds to  $\delta_j$ . Error of solution over the fine  $hp$ -FEM mesh is depending on the relative  $hp$ -FEM error plus the total  $hp$ -FEM error over the coarse mesh plus the accuracy of the proper deme in  $hp$ -HGS tree.

In order to decrease the computational complexity of the algorithm we try to balance the components of the error giving by the right side of the formula (6). Two first components vanishes when  $h \rightarrow 0$  and  $p \rightarrow \infty$  so the third component  $L |g - \hat{d}|$  dominates asymptotically. We perform the  $hp$  adaptation of the FEM solution of the direct problem while the quantity  $\frac{err_{FEM}}{\delta_j}$  is greater then the assumed *Ratio*, which stands for the parameter of the  $hp$ -HGS strategy and corresponds to the Lipschitz constant  $L$  of the energy functional.

Because all  $hp$ -HGS demes work asynchronously on the base of the similar roles, it is enough to present the pseudo-code for the one deme  $P$  of the arbitrary  $j$ -th level in order to define the whole  $hp$ -HGS strategy (see Algorithm 2). The *global\_stop\_condition* appears if the satisfactory set of solutions is found by leaves while the *branch\_stop\_condition* function detects the lack of progress in the evolution of the particular deme. We refer to [5], [10], [9] for more details.

## 4 Asymptotic behavior

The main goal of the asymptotic analysis presented below is to compare  $hp$ -HGS with two other strategies of solving inverse problem (5). The first strategy is the coupling of HGS with the same SGA engines in each branch as in  $hp$ -HGS, but with the fitness function  $f_m$  computed as in  $hp$ -HGS leaves (e.g. by solving the direct problem with the maximum accuracy) and then induced to all branches of lower order. Notice, that such induction is well defined because of the nested HGS encoding (all phenotypes in branches

of the order  $j$  are also phenotypes in branches of the  $j + 1$  order). The second strategy is the single population SGA with the same fitness  $f_m$  as previously. The size of the SGA population ensures the same initial local coverage of the admissible domain  $\mathcal{D}$  by the SGA individuals as by individuals of each  $hp$ -HGS leaf.

We will intensively use the theory of the SGA heuristic (genetic operator) and its fixed points developed by Vose [12] as well as the convergence results of SGA sampling measures (see Schaefer [6], Chapter 4). Let us denote by  $\Omega_j$  the SGA genetic universum of binary codes and by  $G_j : \Lambda^{r_j-1} \rightarrow \Lambda^{r_j-1}$  the genetic operator (heuristic) of all branches (SGA demes) of the order  $j$ . It depends only on the number of genotypes  $r_j$ , fitness function  $f_j$  and the genetic operations applied in branches of the order  $j$ . Moreover the unit simplex  $\Lambda^{r_j-1} \subset \mathbb{R}^{r_j}$  stands for the set of frequency vectors of all possible demes of the order  $j$ . We assume that each genetic operator  $G_j$  has the unique fixed point  $z_j$  in  $\Lambda^{r_j-1}$  that represent the limit population (i.e. the infinite cardinality population after the infinite number of genetic epochs). Moreover  $z_j$  stands for the global attractor of  $G_j$  on  $\Lambda^{r_j-1}$  (i.e.  $\forall x \in \Lambda^{r_j-1} \lim_{t \rightarrow +\infty} (G_j)^t(x) = z_j$ ).

Each deme  $x \in \Lambda^{r_j-1}$  of the order  $j$  may induce the probabilistic measure  $\Theta(x)$  on  $\mathcal{D}$  given by the formula  $\Theta(x)(A) = \sum_{i: \text{code}(i) \in A} x_i$  where  $A \subset \mathcal{D}$  is an arbitrary measurable set (see Section 4.1.2 in [6] for details). Let denote  $\Theta(z_j)$  by  $\Theta_j$  for the sake of simplicity. Moreover we denote by  $\mu_j$  the cardinality of an arbitrary deme of the order  $j$  in  $hp$ -HGS and HGS. We assume that SGA governing the evolution of the  $hp$ -HGS branches of  $j$ -th order  $j \in \{1, \dots, m\}$  are *well tuned* (see Schaefer [6], Definition 4.63). The analogous assumptions are made for the strategy in which the fitness  $f_m$  is implemented in all HGS branches and for the single population SGA.

The course of verifying the asymptotic guarantee of success for  $hp$ -HGS exactly follows the method introduced by Kołodziej [4] and Schaefer [7] for HGS. It consists in proving the compliance of the limit sampling measure of analyzed strategy with the limit sampling measure of the genetic algorithm for which the asymptotic guarantee of success is already checked. The single population SGA with the  $f_m$  fitness was selected as the reference algorithm that posses the desirable feature (see Vose [12] and Schaefer [6]). Another words, we are only obligated to prove, that both  $hp$ -HGS and SGA can found the same set of local minimizers.

**Theorem 4.1.** *Let  $t_0$  be the number of genetic epochs after which  $hp$ -HGS has  $b$  branches of maximal degree  $m$  and no new branches are sprouted. Then*

$$\forall \epsilon > 0, \forall \eta > 0, \exists N \in \mathbb{N}, \exists W(N) > t_0,$$

so that for the arbitrary measurable set  $A \subset \mathcal{D}$

$$\forall \mu_m > N, \forall t > W(N) \quad Pr\{|\chi_{b, \mu_m}^t(A) - \Theta_m(A)| < \epsilon\} > 1 - \eta,$$

where

$$\chi_{b, \mu_m}^t = \frac{1}{b} (\Theta(p_{1, \mu_m}^t) + \dots + \Theta(p_{b, \mu_m}^t))$$

stands for the mean sampling measure on  $\mathcal{D}$ , induced by all  $hp$ -HGS leafs. □

*Sketch of the proof:* The method of verifying above thesis is analogous to the proof of the Theorem 1 in [7]. First we can observed, that for  $b = 1$  it follows immediately from the Theorem 2 in [8].

Now let  $b > 1$  and  $\{p_{1,\mu_m}^t, \dots, p_{b,\mu_m}^t\}$  be a frequency vector of populations evolving in demes of level  $m$  in  $hp$ -HGS tree after  $t > t_0$  metaepoch. Let us fix  $\epsilon > 0$  and  $\eta > 0$ . Because this theorem is true for  $b = 1$  and operator  $G_m$  has the unique fixed point  $z_m$ , then for every deme we have

$$\Pr\{|\Theta(p_{l,\mu_m}^t)(A) - \Theta_m(A)| < \epsilon\} > 1 - \eta$$

where  $N_l \in \mathbb{N}$  and  $W(N_l) \in \mathbb{N}$ ,  $l = 1, \dots, b$ . Because  $\forall l \Theta_{p_{l,\mu_m}^t}(A) < +\infty$  there exist  $l_0$  such that

$$|\Theta_{p_{l_0,\mu_m}^t}(A) - \Theta_m(A)| \geq |\Theta_{p_{l,\mu_m}^t}(A) - \Theta_m(A)|. \quad (7)$$

Let  $l \neq l_0$ ,  $N = \max\{N_1, \dots, N_b\}$ ,  $W(N) = \max\{W(N_1), \dots, W(N_b)\}$ ,  $\mu_m > N$  and  $t > W(N)$ , then

$$\begin{aligned} |\Theta_{p_{l_0,\mu_m}^t}(A) - \Theta_m(A)| &\geq \frac{1}{b} \sum_{l=1}^b |\Theta_{p_{l,\mu_m}^t}(A) - \Theta_m(A)| \geq \\ &\geq \frac{1}{b} \sum_{l=1}^b |\Theta_{p_{l,\mu_m}^t}(A) - b\Theta_m(A)| = \frac{1}{b} b |\chi_{b,\mu_m}^t(A) - \Theta_m(A)| = \\ &= |\chi_{b,\mu_m}^t(A) - \Theta_m(A)|. \end{aligned}$$

Finally, applying (7) for  $\Theta_{p_{l_0,\mu_m}^t}$  we have

$$\Pr\{|\chi_{b,\mu_m}^t(A) - \Theta_m(A)| < \epsilon\} \geq \Pr\{|\Theta_{p_{l_0,\mu_m}^t}(A) - \Theta_m(A)| < \epsilon\} > 1 - \eta$$

which completes the proof.  $\square$

Next we briefly report some others advantageous asymptotic features of  $hp$ -HGS that are presented and proved in detail in [9]. The Corollary 1 in [9] delivers the formula for the computational cost of the single genetic epoch in all  $hp$ -HGS branches

$$\mu_1 a_1 + \sum_{j=2}^m \# \Omega_{j-1} \kappa^j \mu_j a_j \quad (8)$$

where  $a_j$  stands for the average cost of solving direct problem for individuals of the  $hp$ -HGS deme of  $j$ -th order. It is obvious, that  $a_j > a_i$  if  $j > i$ . In particular, when we assume the particular linear regression of the inverse problem error  $\delta_j$  the average cost  $a_j$  is  $\mathcal{O}\left((\theta(j-1) + \beta)^{3\gamma}\right)$  where the constants  $\theta > 0$ ,  $\beta \geq 0$  and  $\gamma > 1$  depend on the inverse problem under consideration (see Corollary 4 in [9]). For typical 3D problems associated with the linear elasticity  $\gamma = 3$  so the mean computational cost grows nine degree for each level in the  $hp$ -HGS tree. The coefficient  $\kappa^j < 1$  stands for the expected ratio of alive branches of the order  $j$ . The computational cost of the single HGS epoch applied to the same inverse problem

$$\mu_1 a_m + \sum_{j=2}^m \# \Omega_{j-1} \kappa^j \mu_j a_m \quad (9)$$

is much greater than the one given by the formula (8). Both HGS and  $hp$ -HGS costs are less than the single deme SGA cost that may be approximated by  $\# \Omega_{m-1} \mu_m a_m$  (see Corollaries 2 and 3 in [9]).

## 5 Simple computational example

In order to illustrate the impact of the twin adaptive strategies in solving parametric inverse problems we briefly quote the simple computational example related to the modeling of the Step-and-flash Imprint Lithography (SFIL). This modern process is a patterning by utilizing photopolymerization to replicate the topography of the template into the substrate. The main goal of this example was to find the proper value of the thermal expansion coefficient (CTE) parameter of the pattern body. The other parameters, like the Young modulus and Poisson ratio were assumed to be equal to  $10^9$  and 0.3 respectively. The initial value of the CTE parameter was set  $-0.4$ . We refer to [11] for more detail description of this experiment.

Because the  $hp$ -HGS algorithm is not fully implemented yet, we used the Hook-Jeeves method as the optimization procedure and  $hp$ -FEM for solving direct problem. The above strategy needs three  $hp$ -FEM calls for each Hook-Jeeves iteration. The rule of accuracy balancing similar to described in section 3 was utilized. The results was compared with the algorithm that solves the same inverse problem with the constant, maximum accuracy of  $hp$ -FEM.

In the first case we take the initial  $hp$ -FEM mesh with 15% relative error of the FEM solution. When the inverse problem was solved up to this accuracy, one step of  $hp$ -adaptation was executed. We obtain the second  $hp$ -FEM mesh with 8% accuracy, and then the third one with 5% accuracy. We needed to perform 43 iterations of algorithm on the first  $hp$ -FEM mesh, 39 on the second mesh and 35 on the third. The time of one iteration equals 0.1s, 1s, 10s respectively. The total execution time was  $43 \times 3 \times 0.1 + 39 \times 3 \times 1 + 35 \times 3 \times 10 = 1179$  seconds. In the second case we used the finest  $hp$ -FEM mesh with the 5% accuracy of direct problem solution and we needed 91 iterations to get the same value of CTE. The total execution time equals  $91 \times 3 \times 10 = 2730$  seconds in this case.

## 6 Conclusions

The proposed  $hp$ -HGS seems to be the advantageous strategy for solving parametric inverse problems for which the energy functional can be defined. Such problems are formulated as the global optimization ones (or optimal control ones) where the objective takes a value of discrepancy between the real energy of the system and its approximated value computed by the algorithm. This class of problems includes important cases of heat flow in solid bodies, fluid flow in porous media and the linear elasticity problems.

In spite of the lower computational cost  $hp$ -HGS offers the same possibility of finding global and local minimizers as HGS and the single population SGA (see Theorem 4.1). The computational cost reduction is resulted from the proper balance between the accuracy of  $hp$ -FEM used for solving direct problem and the accuracy of solving optimization problem.

## Acknowledgements

The financial support of Polish MNiSW grant 3 T08B 055 29 is gratefully acknowledged.

## Bibliography

- [1] Ciarlet P.; *The Finite Element Method for Elliptic Problems*. Society for Industrial & Applied, 2002.
- [2] Demkowicz L.: *Computing with hp-Adaptive Finite Elements. Volume I. One- and Two- Dimensional Elliptic and Maxwell Problems*, Chapman & Hall / CRC Applied Mathematics and Nonlinear Science, 2006.
- [3] Guus C., Boender E., Romeijn E., H.; Stochastic Methods. in Horst R., Pardalos P., M.; *Handbook of Global Optimization*. Kluwer 1995.
- [4] Kołodziej J.; *Hierarchical Strategies of the Genetic Global Optimization*. PhD Thesis, Jagiellonian University, Faculty of Mathematics and Informatics, Kraków 2003, Poland (in Polish).
- [5] Schaefer R., Barabasz B., Paszyński M.; Twin adaptive scheme for solving inverse problems. *Proc. of Conf. on Evolutionary Algorithms and Global Optimization KAEiOG 2007*, 2007, pp. 241–249.
- [6] Schaefer R. (with the chapter 6 written by Telega H.); *Foundation of Genetic Global Optimization*, Studies in Computational Intelligence Series 74, Springer 2007.
- [7] Schaefer R., Kołodziej J.; Genetic search reinforced by the population hierarchy. in De Jong K. A., Poli R., Rowe J. E. eds. *Foundations of Genetic Algorithms 7*, Morgan Kaufman Publisher 2003, pp. 383–399.
- [8] Schaefer R., Jabłoński Z., J.; How to Gain More Information from the Evolving Population? Chapter in: Arabas J. ed. *Evolutionary Computation and Global Optimization*, Warsaw Technical University Press 2002, pp. 21–33.
- [9] Schaefer R., Barabasz B.; Asymptotic behavior of *hp*-HGS (hp-adaptive Finite Element Method coupled with the Hierarchic Genetic Strategy) by solving inverse problems. Accepted to ICCS 2008.
- [10] Paszyński M., Barabasz B., Schaefer R.; Efficient adaptive strategy for solving inverse problems. LNCS Vol. 4487, Springer 2007, pp. 342–349.
- [11] Paszyński M., Szeliga D., Barabasz B., Macioł P.; An algorithm for relating convergence ratios of inverse and direct problem solutions by means of the self-adaptive hp finite element method. 17-th International Conference on Computer methods in Mechanics, June 19-22, 2007, Spała, Poland.
- [12] Vose M., D.; *The Simple Genetic Algorithm*, MIT Press, 1999.