Solving inverse problems by multi-deme hierarchic genetic strategy

R. Schaefer[1], B. Barabasz[2] M. Paszyński[1] S. Migórski[3]

[1]Computer Science Department
[2]Department of Modeling and Information Technology
AGH University of Science and Technology
[3]Institute of Computer Science
Jagiellonian University
Kraków, Poland

Agenda

- Direct and inverse problems
- Basic assumptions of *hp*–HGS
- *hp*–HGS definition
- Lipschitz continuity of the energy functional
- Asymptotic features of *hp*–HGS
- Computational test – heat transfer in the L-shape domain
- Conclusions

Direct problem with energy functional

- The direct problem may be formulated as follows:
  Find $u \in V$ such that

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

  where V is the proper Sobolev space.

- The form b and the functional l depend on the modeled physical phenomena and on its parameter $d \in \mathscr{D}$, where $\mathscr{D}$ is a regular compact in $\mathbb{R}^N$, $N < +\infty$.

- If b is bilinear, symmetric and positively defined, then the energy functional can be defined by

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

hp–adaptive Finite Element Method (*hp*–FEM) I

- The direct problem may be approximated using *hp*–FEM by the following one: Find $u_{h,p} \in V_{h,p} \subset V$ so that

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

- This method creates the h, p-indexed sequence of nested meshes, used for solving finite dimensional problems with the increasing accuracy; h is related to the size of mesh elements, p stands for the polynomial order of approximation.
- Reference: Demkowicz L., Kurtz J., Pardo P., Paszyński M., Rachowicz W., Zdunek A.: Computing with *hp*-Adaptive Finite Elements. Volume II (2007)

hp–adaptive Finite Element Method (*hp*–FEM) II

- The next-step *fine* mesh is constructed by breaking all elements and increasing the order of a polynomial. Both *coarse*- and *fine*-mesh solutions $u_{h,p} \in V_{h,p}$ and $u_{\frac{h}{2},p+1} \in V_{\frac{h}{2},p+1}$ are computed. Both solutions depend also on the parameter $d \in \mathscr{D}$.

- The resulted next-step mesh is obtained on the base of the relative *hp*–FEM error analysis

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

where $\|\cdot\|_E$ stands for the energy norm on the space V. The modifications performed in the fine mesh remain only in elements where its value is large.

Global optimization problem

- We assume that energy $J(\hat{d}) = E(\hat{d}; u)$ of the exact solution $u \in V$ is known and we are looking for the unknown parameter $\hat{d}$.
- Assuming that the proper convergence conditions for (*hp*–FEM) are satisfied the inverse problem can be formulated as follows:

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

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

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

The idea of Hierarchic Genetic Strategy (HGS)

- Tree-structured (depth $< m$), dynamically changing set of dependent demes.
- Low-order demes perform the chaotic low accuracy search finding the promising regions to *sprout* the child-demes.
- Leafs perform the most accurate and local search in the optimization landscape.
- Sprouting is performed conditionally, if there is the room for the new deme among the existing child-demes.
- Basic references: Kołodziej (PhD Thesis 2003), Schaefer and Kołodziej (FOGA Proc. 2003)

Basic assumptions of *hp*–HGS I

- The maximum diameter δ_j of the phenotype grid in $\mathscr{D}$ associated with the demes of the order j determines the search accuracy at this level in *hp*–HGS tree. Of course $\delta_1 > \ldots > \delta_m$.
- 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

$$e_{h,p}(g) = \left| J_{h,p}(g) - J(\hat{d}) \right| ,$$

where $J(\hat{d})$ is the known, real energy and g represents the parameter decoded from the genotype i appears in deme of the level j in *hp*–HGS tree.

Basic assumptions of *hp*–HGS II

- The adaptation of errors is made using the formula

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

  where $e_{\frac{h}{2},p+1}(g) = \left|J_{\frac{h}{2},p+1}(g) - J(\hat{d})\right|$, L stands for the Lipschitz constants of the functionals J and $\left|g - \hat{d}\right|$, the error of the inverse problem solution that corresponds to δ_j.

- First two components vanishes when $h \to 0$ and $p \to \infty$, the third one $L\left|g - \hat{d}\right|$ 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 parameter *Ratio* related to Lipschitz constants.

hp-HGS definition

if ($j = 1$) **then**
  initialize the root deme;
end if
$t \leftarrow 0$;
repeat
  if (*global_stop_condition* received) **then**
    STOP;
  end if
  for ($i \in P^t$) **do**
    solve the direct problem for $g = code(i)$ on the coarse and fine FEM meshes;
    compute $err_{FEM}(g)$;
    while ($err_{FEM}(g) > Ratio * \delta_j$) **do**
      execute one step of *hp* adaptivity;
      solve the problem on the new coarse and fine FEM meshes;
      compute $err_{FEM}(g)$;
    end while;
    compute fitness $f_j(i)$ using the FEM mesh finally established;
  end for
  if ($j > 1$) **then**
    compute the phenotypes' average and send it to the parental deme;
    if (*branch_stop_condition*(P^t)) **then**
      STOP;
    end if
  end if
  if ((($t \bmod K) = 0) \wedge (j < m)$) **then**
    distinguish the best fitted individual x from deme P^t;
    if ($\neg$ *children_comparison*(x)) **then**
      sprout;
    end if
  end if
  perform proportional selection, obtaining multiset of parents;
  perform SGA genetic operations on the multiset of parents;
  $t \leftarrow t + 1$;
until (*false*)

Lipschitz continuity of the energy functional I

The Lipschitz continuity of the energy functional can be verified for the linear heat conduction problem:

$$V = \{v \in H^1(\Sigma),\ tr(v) = 0 \,\mathrm{on}\, \Gamma_D\}$$

where $\Sigma \subset \mathbb{R}^n$ is the bounded compact with the Lipschitz boundary $\Gamma = \Gamma_D \cup \Gamma_N$, $\int_{\Gamma_D} dS > 0$ and $\int_{\Gamma_D \cap \Gamma_N} dS = 0$.

$$Cond = \{r \in L^\infty(\Sigma); 0 < k_0 \leq r(x) \leq K < +\infty\}$$

$$E(k; u(k)) = \frac{1}{2} b(k; u(k), u(k)) - l(u(k)),\ \ k \in Cond$$

where $u(k) \in V$ satisfies

$$b(k; u(k), v) = l(v)\ \ \forall v \in V$$

Lipschitz continuity of the energy functional II

For the heat conduction problem

$$b(k; u(k), v) = \int_{\Sigma} k \nabla u(k) \cdot \nabla v \, dx, \;\; l(v) = \int_{\Sigma} Qv \, dx + \int_{\Gamma_N} \chi \, v \, dS$$

where $Q \in L^2(\Sigma)$, $\chi \in H^1(\Gamma)$.

Theorem

$$\exists L > 0; \forall k_1, k_2 \in Cond$$

$$|E(k_1; u(k_1)) - E(k_2; u(k_2))| \leq L \|k_1 - k_2\|_{L^\infty}$$

Lipschitz continuity of the energy functional III

- The Lipschitz constant L can be evaluated if the continuity constant of the trace operator $tr : H^1(\Sigma) \longrightarrow L^2(\Gamma)$ is known for the particular domain Σ.
- The similar result can be obtained for the linear elasticity problem. The Lipschitz constant may be exactly computed if both: the continuity constant of the trace operator and the Korn constant are well known in this case.

Asymptotic features of *hp*–HGS – preliminaries I

- Let us denote by $G_j : \Lambda^{r_j-1} \to \Lambda^{r_j-1}$ the genetic operator (heuristic) of all branches (SGA demes) of the order j, where $\Lambda^{r_j-1} \subset \mathbb{R}^{r_j}$ is 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 Λ^{r_j-1} that represent the limit population (i.e. the infinite cardinality population after the infinite number of genetic epochs).
- Moreover, we assume, that z_j stands for the global attractor of G_j on Λ^{r_j-1} (i.e. $\forall x \in \Lambda^{r_j-1}\ \lim_{t\to+\infty} (G_j)^t(x) = z_j$).

Asymptotic features of *hp*–HGS – preliminaries II

- Each deme $x \in \Lambda^{r_j - 1}$ of the order j may induce the probabilistic measure $\Theta(x) \in \mathscr{M}(\mathscr{D})$ on $\mathscr{D}$ given by the formula

$$\Theta(x)(A) = \sum_{i;code(i) \in A} x_i$$

where $A \subset \mathscr{D}$ is an arbitrary measurable set and $\mathscr{M}(\mathscr{D})$ denotes all probabilistic measures over the admissible set.

- Let denote $\Theta(z_j)$ by Θ_j for the sake of simplicity.
- Moreover denote by μ_j the cardinality of an arbitrary deme of the order j in *hp*–HGS and HGS.

Asymptotic features of *hp*–HGS – preliminaries III

- Each deme $x \in \Lambda^{r_j-1}$ induces the probabilistic measure on $\mathscr{D}$ given by the density $\rho_x \in L^p(\mathscr{D}), p \geq 1$.
- $\Psi_j : \Lambda^{r_j-1} \to \mathscr{M}(\mathscr{D})$ such that $\Psi_j(x)(A) = \int_A \rho_x(\xi) d\xi$ for each measurable set $A \subset \mathscr{D}$.
- The densities ρ_x, $x \in \Lambda^{r_j-1}$ are piecewise constant on some subsets surrounding the phenotypes induced by the encoding of j-th order (e.g. on the Voronoi neighborhoods of phenotypes).
- Let us denote $\psi_j = \Psi_j(z_j)$ for the sake of simplicity.

Asymptotic features of *hp*–HGS – preliminaries IV

- We assume that SGA governing the evolution of the *hp*–HGS branches of j-th order $j \in \{1, \ldots, m\}$ are *well tunned* i.e. the densities ρ_j dominates on some closed sets $C^j \subset \mathscr{D}$ with the strictly positive Lesbegue measure (not necessary connected).
- We assume that $C^1 \supset C^2 \supset, \ldots, \supset C^m$.
- 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.

Asymptotic features of *hp*–HGS – results I

Theorem

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

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

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

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

where

$$\chi^t_{b,\mu_m} = \frac{1}{b}\left(\Theta(p^t_{1,\mu_m}) + , \ldots , + \Theta(p^t_{b,\mu_m})\right)$$

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

Asymptotic features of *hp*–HGS – results II

Theorem

The hp–HGS deme of the order $j_0 = 2, \ldots, m$ survives in the hp–HGS tree with the probability not greater than κ^{j_0} given by the formula:

$$\kappa^{j_0} = \psi_1(C^1) \prod_{j=2}^{j_0} \frac{\psi_j(C^j)}{\psi_j(C^{j-1})}.$$

The proofs of these theorems are based on the fixed point theory of SGA heuristics (Vose(1999), Vose and Nix(1992)) and the analysis of the HGS efficiency (Schaefer, Kolodziej(2002)).

Asymptotic features of *hp*–HGS – results III

- The computational cost of the single genetic epoch for *hp*–HGS, HGS and SGA may be approximated by:

$$\mu_1 \, a_1 + \sum_{j=2}^{m} \sharp\Omega_{j-1} \, \kappa^j \, \mu_j \, a_j,$$

$$\mu_1 \, a_m + \sum_{j=2}^{m} \sharp\Omega_{j-1} \, \kappa^j \, \mu_j \, a_m,$$

$$\mu_1 \, a_m + \sum_{j=2}^{m} \sharp\Omega_{j-1} \, \mu_j \, a_m,$$

  where a_j stands for the average cost of solving the direct problem for individuals of the *hp*–HGS deme of j-th order and $a_j > a_i$ if $j > i$. For the properly profiled algorithm κ^j is much less than 1. $\sharp\Omega_{j-1}$ denotes the cardinality of the space of binary codes for the branch of $j-1$-th order.

- $cost_{hp-HGS} \le cost_{HGS} \le cost_{SGA}$.

Asymptotic features of *hp*–HGS – results IV

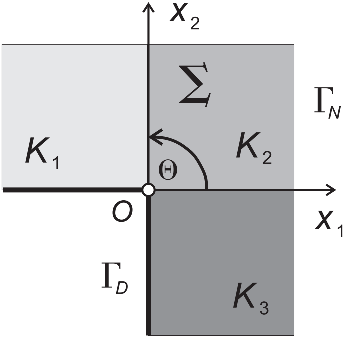
- When we assume the linear regression of $\{\delta_j\}, j = 1, 2, \ldots, m$ and the exponential regression of the *hp*-FEM error then the average cost a_j is

$$\mathscr{O}\left(\left(\theta\left(j-1\right)+\beta\right)^{3\gamma}\right)$$

  where the constants $\theta > 0$, $\beta \geq 0$ and $\gamma > 1$ depend on the inverse problem under consideration.

- 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!

L-shape domain direct problem I



L-shape domain direct problem II

The strong formulation of the direct problem consists in finding the temperature distribution $u \in C^2\left(\mathbb{R}^2\right)$ such that

$$\nabla \cdot k\nabla u = 0 \text{ in } \Sigma \subset \mathbb{R}^2$$

with boundary conditions

$$u = 0 \text{ on } \Gamma_D, \ \frac{\partial u}{\partial n} = \chi = r^{\frac{3}{2}} sin^{\frac{3}{2}}\left(\theta + \frac{\pi}{2}\right) \text{ on } \Gamma_N$$

L-shape domain problem III

Weak formulation of the direct problem:
Find

$$u \in V = \{v \in H^1(\Sigma); tr(v) = 0 \text{ on } \Gamma_D\}$$

such that:

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

$$b(d; u, v) = \int_\Sigma k(d) \nabla u \cdot \nabla v \, dx, \;\; l(v) = \int_{\Gamma_N} \chi \, v \, dS$$

where $d = (K_1, K_2, K_3) \in \mathscr{D} = [0,3]^3 \subset \mathbb{R}^3$ is the vector of parameters. $k(d)$ takes the constant values K_1, K_2, K_3 over three parts of the domain.

L-shape domain problem IV

- The energy of the exact solution, was computed as $J_{h,p}(\hat{d})$, where $\hat{d} = (0.1, 1.5, 2.9)$ are the assumed values of the heat transfer coefficients.
- In order to evaluate the exact energy $J(\hat{d})$ by $J_{h,p}(\hat{d})$ the L-shape domain problem was computed once with the very high relative accuracy 10^{-5} by using the self-adaptive *hp*-FEM code.
- Because both boundary conditions are symmetric with respect to the domain geometry (e.g. with respect to the axis $\theta = \pi/4$) then the parameter vector $\hat{d}' = (2.9, 1.5, 0.1)$ gives the same energy as $\hat{d}$, so the inverse problem has two solutions in this case.

Test I

Standard deviations of phenotypes for various mutation rates tested for all levels of the *hp*–HGS tree

Mutation rate	level 1	level 2	level 3
0.5	1.54	1.51	1.51
0.2	1.29	1.19	1.21
0.1	0.85	0.91	0.83
0.035	0.49	0.55	0.56
0.025	0.4	0.53	0.55
0.01	0.31	0.23	0.25
0.005	0.24	0.18	0.11

Tests II

Final parameters of the *hp*–HGS tree applied by solving the sample inverse problem

	level 1	level 2	level 3
Code length	9	18	27
δ_j	$58 * 10^{-4}$	$11 * 10^{-6}$	$22 * 10^{-9}$
Population size	100	40	10
Crossing rate	0.5	0.5	0.5
Mutation rate	0.2	0.02	0.002

Where δ_j is the search accuracy on each level of the *hp*–HGS tree

Tests IV

	test 1	test 2	test 3
number of leafs	13	6	6
Ratio	$44.7 * 10^7$	$22.3 * 10^7$	$44.7 * 10^6$
accuracy on 3-th level	10	5	1
err_{fit}	22.85%	0.29%	0.06%
min $eucl_1$	0.38	0.01	0.09
min $eucl_2$	0.23	0.06	0.007

Conclusions

- The proposed *hp*–HGS seems to be the advantageous strategy for solving parametric inverse problems for which the energy functional can be defined. This class of problems includes important cases of heat flow in solid bodies, fluid flow in porous media and the linear elasticity problems.
- *Ratio* coefficient determining the accuracy of the *hp*–FEM was decreased in the consecutive tests 1,2 and 3. It was follwed by decreasing fitness error err_{fit} and the minimum distances between phenotypes and both minimizers. The proper redundancy reduction policies decreases the number of leaves in the later tests 2 and 3.
- The idea of *hp*–HGS can be extended to other cases of global optimization problems in continuous domains in which the computational cost of the objective evaluation depends monotonically on its accuracy.