ON MESH REFINEMENT FOR EVOLUTIONARY ALGORITHMS ACCELERATION APPLIED TO LARGE NON-LINEAR CONSTRAINED OPTIMIZATION PROBLEMS

Janusz Orkisz, Maciej Głowacki

Institute for Computational Civil Engineering
Cracow University of Technology
Contents

1. Introduction
2. Acceleration techniques considered
3. Step by step mesh refinement
4. A’posteriori solution error analysis
5. Mesh refinement combined with error analysis
6. Benchmark problems analysis
7. Final remarks
Introduction

Research motivation

A variety of engineering and scientific tasks may be formulated as large, non-linear, constraint optimization problems, e.g.:

- residual stress analysis in railroad rails and vehicle wheels,
- physically based approximation of experimental data.

Their effective solution is crucial for various practical engineering problems.

Solution methods used (convex, and non-convex problems):

deterministic like FDM (Feasible Directions Method) or Penalty Methods and/or probabilistic like AI (e.g. Evolutionary Algorithms)

Research objective

Significant acceleration of the EA applied to large, non-linear constraint optimization problems, where a function (given e.g. by its nodal values) is searched.

The speed-up is based on:

- choice of the most effective combination of the evolutionary operators: selection, crossing-over, mutation,
- use of several new simple speed-up techniques proposed here,
- further development of chosen existing methods.
Typical formulations of these optimization problems

(i) Residual stress analysis

Find minimum of the total complementary energy

$$\min_{\sigma^s} W(\sigma^s - \sigma^R_0) , \quad W(\sigma^s - \sigma^R_0) = \frac{1}{2} \int_V (\sigma^s - \sigma^R_0)^t C (\sigma^s - \sigma^R_0)$$

for self-equilibrated stresses $\sigma^s$ satisfying

- **equilibrium** equations
  $$\partial \sigma^s = 0 \quad \text{in} \quad V$$

- **boundary** conditions
  $$\nu^t \sigma^s = 0 \quad \text{on} \quad \partial V$$

and **yield** constraint

$$\Phi (\sigma^s + \sigma^E) \leq 0 \quad \text{in} \quad V \cup \partial V$$

where

- $W$ - total complementary energy functional corresponding to stresses $\sigma$ in the whole body $V$
- $\sigma^R_0$ - initial self-equilibrated stresses
- $C$ - compliance matrix
Typical formulations of these optimization problems

(ii) Physically based approximation

Find \( \sigma = \sigma(x) \) that yields the stationary value of the functional

\[
\Phi = (1 - \lambda) \Phi^T + \lambda \Phi^E \quad \lambda \in [0,1]
\]

and satisfy

- **equality** constraints (theoretical)
  \[ A(\sigma) = 0 \]

- **inequality** constraints (experimental)
  \[ B(\sigma) \leq e \]

where

\[
\Phi^T = \frac{\Phi^T}{\Phi^T_{\text{REFERENCE}}} \quad \text{- dimensionless theoretical part of the weighted functional } \Phi^T
\]

\[
\Phi^E = \frac{\Phi^E}{\Phi^E_{\text{REFERENCE}}} \quad \text{- dimensionless experimental part of the weighted functional } \Phi^E
\]

\[ \lambda \in [0,1] \quad \text{- weighting factor} \]

\[ e \quad \text{- admissible tolerance} \]
Standard **Evolutionary Algorithm** used

Choice of the **most effective combination** of evolutionary operators

- **Selection** operators:
  - rank
  - tournament

- **Crossing-over** operators:
  - simple
  - arithmetic
  - heuristic

- **Mutation** operators:
  - uniform
  - non-uniform
  - boundary

Initial population generation

Selection

Crossover

Mutation

**BREAK OFF TEST**

YES

NO

END
Acceleration techniques considered

**Newly proposed** simple acceleration techniques:

- **smoothing** of the direct EA solution,
- **balancing** of smoothed EA solution,
- step by step **mesh refinement**, 
- use of **a’ posteriori error analysis** and non-standard **parallel** and **distributed** calculations, combined with **mesh refinement**.

**Development of chosen** **existing** techniques:

- **new** evolutionary **operators** (e. g. gradient mutation, cloning),
- **hybrid** algorithms (EA + deterministic method).
- **distributed** and **parallel** algorithms.
Step by step mesh refinement

**Adaptive** approach

- nodes of **old** cloud
  (with known solution)

- candidates for the **new** cloud

**Generation criterion** of new nodes

\[ e_x > \tau \cdot \overline{e}_{\text{max}}, \quad 0 < \tau < 1 \]

**Break off criterion**

\[ \forall \overline{e}_x < e_{adm} \]

**Use of a’posteriori error analysis** for estimating the quality of nodes distribution and adaptation approach
Step by step mesh refinement

Regular meshes (for benchmark problems)

1D

Initial function values in inserted nodes may be calculated using interpolation spanned over coarse mesh

2D

- nodes of old cloud
- new nodes
**Smoothing by 1D MWLS approximation technique**

Weighted error functional

\[ B = \sum_{i=1}^{n} (u_i - \bar{u}_i)^2 w_i^2 \]

where

\[ w_i^2 = \left( \frac{h_i^2 + \frac{g^4}{h_i^2 + g^2}}{h_i^2 + g^2} \right)^{-p-1} \]

\[ \bar{u}_i = \bar{u}(x_i) = \bar{u} + h_i \bar{u}' + \frac{1}{2} h_i^2 \bar{u}'' + \ldots + \frac{1}{p!} h_i^p \bar{u}^{(p)} + R \]

\[ \bar{u} = \bar{u}(\bar{x}) \]

Minimization conditions:

\[ \frac{\partial B}{\partial \bar{u}} = 0, \quad \frac{\partial B}{\partial \bar{u}'} = 0, \quad \frac{\partial B}{\partial \bar{u}''} = 0, \quad \ldots, \quad \frac{\partial B}{\partial \bar{u}^{(p)}} = 0 \]

\[ \Rightarrow \bar{u}, \ \bar{u}', \ \bar{u}'', \ \ldots, \ \bar{u}^{(p)} \]

Weighting function for various \( g \) values:
A’ posteriori solution error analysis

Solution of optimization problem

\[ u^* \text{ true} \]

\[ \widetilde{u} = \widetilde{u}_{h, p} \text{ rough solution} \]

\[ \bar{u} \text{ improved reference solution - hierarchic approach} \]

new solutions and postprocessing required

Local a’ posteriori error

\[ e_T = \widetilde{u} - u^* \text{ true error} \]

\[ e_E = \widetilde{u} - \bar{u} \text{ estimated error} \]
A’ posteriori solution error analysis

**Global a’ posteriori solution error**

\[ e = \| e_T \| \quad \text{true error} \]
\[ \eta = \| e_E \| \quad \text{estimated error} \]

Error estimation type: **hierarchic** - h, p, smoothed HO

Error norms

- **mean square:**
  \[ e_{\text{MEAN}} = \left( \frac{1}{\Omega} \int_{\Omega} e_T^2 \, d\Omega \right)^{\frac{1}{2}} \]
  \[ \eta_{\text{MEAN}} = \left( \frac{1}{\Omega} \int_{\Omega} e_E^2 \, d\Omega \right)^{\frac{1}{2}} \]

- **maximum:**
  \[ e_{\text{MAX}} = \max_{x \in \Omega} e_T \]
  \[ \eta_{\text{MAX}} = \max_{x \in \Omega} e_E \]

Estimation quality – **effectivity index**

\[ i = 1 + \frac{\| e \| - \| \eta \|}{\| e \|} \]

Fitness function value \( f_{\text{best}} \) control

\[ \xi = \left| \frac{f_{\text{best}}^{(m)} - f_{\text{best}}^{(m-1)}}{f_{\text{best}}^{(m)}} \right| \leq \xi_{\text{admissible}} \]
Generation of reference solutions

In the **Evolutionary Algorithms**

**Approach**
- generation of population of *m* independent solutions (chromosomes),
- weighted averaging of these results over the whole population,
- postprocessing (HO smoothing) of the above discrete solution by means of the MWLS (or PBA) approximation,

**Notice**
- *h*-type and/or *p*-type estimation approach would be also possible.
Use of a’ posteriori error analysis for the EA solution process acceleration

Approach concept

- Find or estimate EA solution error in the whole domain.
- **Intensify calculations** in zones where solution error is larger than elsewhere.

Do this by means of modification of the basic EA operators:

- **selection**, **mutation**, and **crossing-over**.

Modification of the selection operator

Examine various ways of **simultaneous** use of two selection criteria

- **fitness** function value $f$,
- solution error value $e$.  


Modification of the **mutation** operator

Increasing **probability** of mutation:
- in regions where error is **larger** than e.g. $\frac{1}{2}$ mean error,
- following function of the error level.

Modification of the **crossing-over** operator

For heuristic **crossing-over**:
- crossing-over in **direction** of the **best** of three chromosomes:
  mean $M$ and two parents $X, Y$,
e.g. If $F(M) > F(P)$ :

$$Z = a(M - P) + P , \quad P = \sup \{X, Y\}$$
otherwise
  standard procedure

- increasing **value** of the parameter $a$.

Other: increasing **crossing-over probability** for values larger than e.g. $\frac{1}{2}$ mean error.

\[
\begin{align*}
P_M & \quad \varepsilon \\
\mid \varepsilon & \quad \mid \varepsilon
\end{align*}
\]

<table>
<thead>
<tr>
<th>Standard heurisic crossing-over:</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X, Y$ - parents,</td>
</tr>
<tr>
<td>$Z$ - offspring,</td>
</tr>
<tr>
<td>$F$ - fitness function.</td>
</tr>
</tbody>
</table>

If $F(X) > F(Y)$

$$Z = a(X - Y) + X$$

$a \in [0,1]$
Use of mesh refinement with a’posteriori error estimation

Strategy

(i) calculation of solution on a coarse mesh
(ii) smoothing of rough solution
(iii) mesh refinement and the best approximation (or interpolation) of initial function values at inserted nodes
(iv) use of obtained solution as the initial reference solution for a’posteriori error estimation
(v) repetition of the procedure given above until a sufficiently dense mesh is reached
Benchmark tests

(1) **Residual stress** analysis in a bar subject to pure cyclic bending – **exact 1D model**

Problem formulation:

Find \( \min I \)

**Fitness** function: \( \min \sigma \int_0^H \sigma^2 dz \) complementary energy

**constraints:** \( M = \int_0^H \sigma z dz = 0 \) global equilibrium

\[ |\sigma + \sigma^e| \leq \sigma_y \] yield condition for total stresses

where:

\( \sigma_y \) – yield stress
\( \sigma^e \) – elastic solution of the problem
\( E \) – Young modulus

NOTICE:

- **EXACT SOLUTION IS KNOWN**
- **ANY NUMBER** OF DECISION VARIABLES MAY BE CHOSEN (IN 1D OR 2D)
Final discrete 1D formulation of the problem:

Find stresses $\sigma_1, \sigma_2, \ldots, \sigma_n$ satisfying:

$$\min_{\sigma_1, \ldots, \sigma_{n-1}} \left( \sum_{k=1}^{n-1} \sigma_k^2 + \frac{1}{2} \sigma_n^2 \right)$$

minimum of the total complementary energy

$$\sigma_n = - \frac{2}{z_n} \sum_{k=1}^{n-1} \sigma_k z_k$$

global equilibrium equation

$$-1 \leq \frac{\sigma_k}{\sigma_y} - 3 \left[ 1 - \frac{1}{3} \left( \frac{\bar{Z}}{H} \right)^2 \right] \frac{k-1}{n-1} \leq 1$$

condition for total stresses (plastic limit)

$k = 1, \ldots, n$

Data assumed: $\sigma_y = 1$; $\frac{\bar{Z}}{H} = \frac{1}{2}$. 
Benchmark tests

(2) **Discrete 2D model**

**Fitness** function:

Total complementary energy

\[
I = \int_{-b/2}^{b/2} \int_{-H}^H \sigma^2 \, dz \, dy \approx \frac{h^2}{9} \left( \sum_{k=1}^{n} \sigma_k^2 \alpha_k \right), \quad \alpha_j - \text{integration coefficient}
\]

**Constraints:**

bending moment – equilibrium equation

\[
M = \int_{-b/2}^{b/2} \int_{-H}^H \sigma \, z \, dz \, dy \approx \frac{h^2}{9} \left( \sum_{k=1}^{n-1} \sigma_k z_k \alpha_k + \sigma_n z_n \alpha_n \right) = 0 \quad \Rightarrow \quad \sigma_n = \frac{\sum_{k=1}^{n-1} \sigma_k z_k \alpha_k}{z_n \alpha_n}
\]

yield condition constraints

\[-1 \leq \frac{\sigma_k + \sigma_k^e}{\sigma_Y} \leq 1, \quad k = 1, \ldots, n\]

\[
\sigma_k^e = \frac{3 \sigma_Y}{2H} \left[ 1 - \frac{1}{3} \left( \frac{Z}{H} \right)^2 \right] z_k \quad z_k = k \frac{H}{n}
\]

\[
b \times 2H \quad - \text{size of beam cross-section}
\]

\[
\sigma_Y \quad - \text{yield stress}
\]

\[
\sigma_e \quad - \text{purely elastic solution stress}
\]
Benchmark tests

(3-4) Residual stresses analysis in thick-walled cylinder under cyclic pressure

Exact (axially symmetrical) formulation

Find minimum of the total complementary energy

$$\min_{\sigma_r^e, \sigma_i^e, \sigma_z^e} \Pi, \quad \Pi = \frac{1}{2} \int_a^b \left[ (\sigma_r^e - \sigma_i^e)^2 + (\sigma_i^e - \sigma_z^e)^2 + (\sigma_z^e - \sigma_r^e)^2 \right] r \, dr$$

subject to

**equilibrium** equation

$$\frac{\partial \sigma_r^e}{\partial r} + \frac{\sigma_i^e - \sigma_r^e}{r} = 0$$

**boundary** conditions

$$\sigma_{r_{ld}}^e = 0, \quad \sigma_{r_{ip}}^e = 0$$

**incompressibility** equation

$$\sigma_z^e = \nu (\sigma_r^e + \sigma_i^e)$$

**yield** condition

$$\phi(\sigma_r^e, \sigma_i^e, \sigma_z^e, \sigma^E) \leq \sigma_y$$

Benchmark (3) - model 1D

Benchmark (4) - model 2D
Preliminary results for benchmark problems – introduction

Four types of charts for speed-up $S$ determination

\[ S_{e,n} = \frac{n_{REF}}{n} \quad e=\text{const} \]

\[ S_{F,n} = \frac{n_{REF}}{n} \quad F=\text{const} \]

All four above situations together provide the full acceleration characteristics investigated.
Preliminary results for benchmark problems – introduction

Two types of charts for precision increase $P$ determination

$$P_n = \frac{e_{REF}(n)}{e(n)}$$

$$P_t = \frac{e_{REF}(t)}{e(t)}$$
**T1. Motivation example**

Impact of mesh density on the convergence rate and solution precision

**Benchmark 3**
T2. Examination of the possibility of solving large optimization problems using mesh refinement and error estimation: benchmark (2)

\[ \sigma(y, z) \]

THE TRUE SOLUTION
T2. Examination of the possibility of solving large optimization problems using mesh refinement and error estimation: benchmark (2)

The true error map

Convergence of mean solution error – comparison with standard algorithm

Speed-up: about 20 - 400 times
Precision increase: about 150 times
T3. Examination of the possibility of solving large optimization problems using mesh refinement and error estimation: benchmark (3)

1D model

7 nodes

13 nodes

25 nodes

49 nodes

97 nodes

Errors

<table>
<thead>
<tr>
<th>mesh</th>
<th>mean err.</th>
<th>max err.</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>0.0208</td>
<td>0.0408</td>
</tr>
<tr>
<td>13</td>
<td>0.0053</td>
<td>0.0114</td>
</tr>
<tr>
<td>25</td>
<td>0.0016</td>
<td>0.0033</td>
</tr>
<tr>
<td>49</td>
<td>0.0011</td>
<td>0.0023</td>
</tr>
<tr>
<td>97</td>
<td>0.0008</td>
<td>0.0018</td>
</tr>
</tbody>
</table>
T3. Examination of the possibility of solving large optimization problems using mesh refinement and error estimation: benchmark (3)

1D model

Continuation of mesh refinement:

97 → 193 → 385 → 769 → 1537 → 3073 nodes

Convergence of mean solution error – comparison with standard algorithm

<table>
<thead>
<tr>
<th>Nodes</th>
<th>193 nodes</th>
<th>3073 nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean err.</td>
<td>0.000905</td>
<td>0.000903</td>
</tr>
<tr>
<td>Max err.</td>
<td>0.001984</td>
<td>0.001980</td>
</tr>
</tbody>
</table>

Speed-up: about 120 times
Precision increase: about 90 times
T4. Examination of the possibility of solving large optimization problems using mesh refinement and error estimation: benchmark (4)

2D model

Solution in cross-section

Discretization strategy

- nodes of **old** mesh
- **new** nodes
T4. Examination of the possibility of solving large optimization problems using mesh refinement and error estimation: benchmark (4)

**2D model**

**solution in cross-section** for 2145 nodes and mesh 65 x 33

**convergence of mean solution error** – comparison with standard algorithm

**Speed-up**: about 150 - 300 times

**Precision increase**: about 100 times
T5. Comparison of the **exact** and **estimated** stress errors for benchmark (2) after 3000 iterations

EXACT ERROR

ESTIMATED ERROR

reference solution obtained by averaging 12 independent chromosomes

**DIFFERENCE BETWEEN THE EXACT AND ESTIMATED ERROR IN BEAM CROSS-SECTION**

Mean values:

<table>
<thead>
<tr>
<th></th>
<th>$e_i$</th>
<th>$e_i / e_{exact}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>exact</td>
<td>0.338</td>
<td>1</td>
</tr>
<tr>
<td>estimated</td>
<td>0.326</td>
<td>0.96</td>
</tr>
<tr>
<td>difference</td>
<td>0.095</td>
<td>0.28</td>
</tr>
</tbody>
</table>

**effectivity index**

$$i = 1 + \frac{\|e\|}{\|\nu\|} = 1.04$$
Final remarks

Summary

- Many scientific and technical tasks may be expressed in terms of non-linear, constraint optimization problems. In a wide class of such problems the objective is to find an unknown function, mostly in a discrete form.

- Optimization problems may be solved by means of either deterministic or probabilistic methods. The first ones are very effective when dealing with the convex problems as opposed to usually slowly convergent probabilistic e.g. most AI methods, especially for large optimization problems. However, their efficiency does not change much for non-convex methods as opposed to the deterministic ones.

- The objective of this whole research, therefore, is to develop means of essential acceleration of the Evolutionary Algorithms, one of the AI methods. Particular attention is paid here to use step by step mesh refinement and use our knowledge about estimated solution error to the EA acceleration. These concepts are tested on various, carefully selected benchmark problems, using the true solution error, replaced later on by the estimated errors. Solutions based on series of more and more dense meshes are used as the reference ones then.
Summary

- Very preliminary results of these tests are encouraging. For use of step by step mesh refinement together with a’posteriori error analysis, the overall speed-up factor about 150 times was reached.

- When well designed, the EA speed-up may provide solution much faster than the standard EA algorithm. Moreover, as opposed to the standard approach such solution may also be obtained for large optimization problems.

Further research planned

- continuation of various efforts oriented towards increasing the EA efficiency, including testing further concepts, and combining all types of acceleration together for their simultaneous use;

- analysis of further benchmarks;

- application to residual stress analysis in railroad rails and vehicle wheels;

- analysis of large, non-linear, constrained optimization problems (convex and non-convex) resulting from the physically based approximation applied to experimental measurements.
THANK YOU VERY MUCH FOR ATTENTION
Generacja węzłów

$p = \log \frac{r}{r_{\min}}$

siatka 1D

<table>
<thead>
<tr>
<th>$p$</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>2</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>1</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>

siatka 2D

\[
p = \begin{cases} 
\min (p_x, p_y) & \text{for } p_x \neq p_y \\
 p_x + 0.5 & \text{for } p_x = p_y 
\end{cases}
\]

<table>
<thead>
<tr>
<th>$p_x$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>0</th>
<th>1</th>
<th>0</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>2</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>1</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>

$p$ = [0, 0.5, 1, 1.5, 2, 2.5, 3, 3.5]
Generacja siatki metodą Liszki

- Zdefiniowanie lokalnej gęstości siatki regularnej

\[ \rho^{-1} \equiv \log_2 \frac{r}{r_{\text{ref}}} = \log_2 \frac{r}{r_{\text{min}}} + \log_2 \frac{r_{\text{min}}}{r_{\text{ref}}} \]

\[ p = \log_2 \frac{r}{r_{\text{min}}}; \quad \rho_{\text{max}}^{-1} = \log_2 \frac{r_{\text{min}}}{r_{\text{ref}}} \]

and \( r_{\text{ref}} \leq r_{\text{min}} \)

jest założonym najmniejszym modułem siatki regularnej

stąd

\[ \rho^{-1} - \rho_{\text{max}}^{-1} = p \]

- Przypisanie żądanej gęstości \( \bar{p} \)

- Wybór węzłów z siatki na podstawie kryterium

\[ \rho^{-1} - \rho_{\text{max}}^{-1} \geq \bar{p} \]
Przykład 1D

\[ \rho^{-1} \geq \bar{p} \]

The diagram illustrates a mesh of nodes and points with numbers characterizing local reversed mesh density factors. The mesh is marked with shapes representing different weighting factors p for the 1D mesh.
Definicja gęstości siatki nieregularnej

- Gęstość siatki
  - gęstość siatki $\rho_i$ w dowolnym węźle $i$ to
  $$\rho_i^{-1} = \rho_{\text{max}}^{-1} + p_i,$$
  gdzie
  $$p_i = \log_2 \left( \frac{\Omega_i}{\Omega_{\text{min}}} k \right)^{1/2}$$

  - gęstość siatki $\rho$ w dowolnym punkcie $x$ obszaru jest obliczania przez aproksymację MWLS
  $$\rho = \sum_{\text{star}} \rho_i \Phi_i,$$
  gdzie $\Phi_i = \Phi_i(x)$ są pseudo – funkcjami kształtu MLWS