Augmented Downhill Simplex a Modified Heuristic Optimization Method

Mohsen Jalaeian-F

Department of Electrical Engineering, Center of Excellence on Soft Computing and Intelligent Information Processing (SCIIP), Ferdowsi University of Mashhad, Mashhad, Iran

Abstract

Augmented Downhill Simplex Method (ADSM) is introduced here, that is a heuristic combination of Downhill Simplex Method (DSM) with Random Search algorithm. In fact, DSM is an interpretable nonlinear local optimization method. However, it is a local exploitation algorithm; so, it can be trapped in a local minimum. In contrast, random search is a global exploration, but less efficient. Here, random search is considered as a global exploration operator in combination with DSM as a local exploitation method. Thus, presented algorithm is a derivative-free, fast, simple and nonlinear optimization method that is easy to be implemented numerically. Efficiency and reliability of the presented algorithm are compared with several other optimization methods, namely traditional downhill simplex, random search and steepest descent. Simulations verify the merits of the proposed method.

Keywords: Augmented Downhill Simplex Method (ADSM); Downhill Simplex; global optimization; global exploration

1. Introduction

Downhill Simplex method (also called Nelder-Mead or Amoeba) is introduced by Nelder and Mead in 1996 [1] which is a useful heuristic optimization method without any assumption on the cost function to minimize. In particular, the cost function does not need to satisfy any condition of differentiability or linearity. So, the most important merits of this method are: 1- it does not have any limitation on cost function (differentiability, linearity or nonlinearity), 2- it is very useful in many-dimensional optimization problems, 3- it is a robust algorithm (derivative-free), 4- it is a fast local optimization method, and 5- it is easy to be implemented numerically. Due to these advantages the DSM is an interesting method and cited by scientists in a wide area of knowledge. This fact is illustrated in Fig. 1, whereas the numbers of publications which use this method are counted versus years (in IEEE, IEEE conferences and Elsevier publishers). However, the only drawback of this method is that it is a local optimizer and can be trapped in a local minimum. The DSM is not to be confused with the Simplex Method [2] that is used for linear programming problems, even though they both use the concept of multidimensional simplex to reach their goal.

During past decades, several modifications on the downhill simplex algorithm are introduced. For instance, DSM is combined with other optimization algorithms such as, Monte-Carlo [3, 4], GP and RLS [5], GA [6-14], differential evolution (DE) [15, 16], finite element [17], pattern search [18], Boltzmann probability [19], K-mean algorithm [20] and simulated annealing (SA) [21-24] to introduce adaptive simplex simulated annealing algorithm (ASSA). Also in the work of Huang et al. multi-start DSM was proposed by running DSM a number of times serially [25, 26]; or parallel (Shuffled Complex Evolution (SCE)) [16, 18, and 27]. In some works, constraints are applied to the position of vertices and Constrained-DSM is made [28-30]. Also, the DSM is applied to a variety of applications such as astronomy [31], self-orientation of directional antennas [32], biosource [33], sensor position optimization [34], energy [35], control [36], earth quake detection [37], motion estimation in video [38], etc.

As mentioned, a good variety of modifications have been done on DSM; all of them improved *reliability* versus the pure DSM. With *"reliability"* we mean the ability for finding the global minima. But unfortunately, two of the most important advantages of this algorithm, which are *"speed of convergence"* and *"ease of implementation"*, are subsided. The set of both these advantages can be called

* Corresponding author. Email: m.jalaeian@yahoo.com

"efficiency". So, we need another proper modification that makes improvement on *reliability* and *efficiency* together.

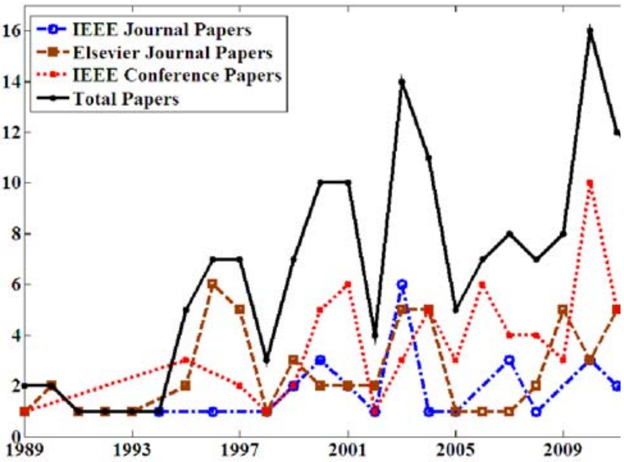

Fig. 1. Number of Publication on Downhill Simplex during years (In IEEE and Elsevier)

In fact, we will to improve the power of finding global minima, and do not damage the speed of convergence or the simplicity of the algorithm. In parallel, nowadays, thanks to the power of numerically computations, we can compute vectors and matrices faster. In results, one can use a vector in a simple way inside DSM computations and it is feasible now to combine some other very simple algorithms with DSM.

This paper is organized as follow: first, the DSM is described, and next, the proposed ADSM is explained. Then, simulations and comparisons are discussed, and the results are summarized in conclusion.

2. Downhill Simplex Algorithm

Downhill Simplex method is a commonly used optimization method to minimize a cost function of n variables. The method relies on the comparison of function values at the n+1 vertices of a general simplex (polytope of dimension n+1). As instance, one dimensional simplex is a line, in two dimensions, the simplex is a triangle and in three dimensions, it is a tetrahedron. The algorithm progresses iteratively while replacing the worst vertex (with the highest cost value) by another projected point. *"The simplex adapts itself to the local landscape, and contracts on to the final minimum."* [1].

The projected point is determined by one of the following operators: reflection, expansion, contraction and shrinking, which produce P_r, P_e and P_c respectively through (1).

$$\begin{aligned} P_r &= (1+\alpha)P_{centroid} - \alpha P_h \\ P_e &= (1+\gamma)P_{centroid} - \gamma P_h \\ P_c &= (1+\beta)P_{centroid} - \beta P_h \end{aligned} \quad (1)$$

In which, P_h is the worst vertex of simplex and $P_{centroid}$ is the centroid of all pointes except P_h. Constants α, γ and β are the reflection, extraction and contraction coefficients respectively. The Downhill Simplex algorithm is explained below:

0- Initialize $0<\alpha<1$, $1<\gamma$ and $-1<\beta<0$
1- Initialize vertexes of the simplex
2- Sort the simplex vertexes on depend on the cost function, and find P_l , P_h , $P_{centroid}$, also compute P_r
3- If $P_l < P_r < P_h$; then $P_r \rightarrow P_h$
4- Else if $P_r < P_l$ (P_r product new minimum), then compute P_e and
a. if $P_e < P_l$ then $P_e \rightarrow P_h$
b. else if $P_e > P_l$ then $P_r \rightarrow P_h$
5- Else if $P_r > P_k$ for all $k \neq h$ then compute P_c and
a. if $P_c < min(P_h, P_r)$ then $P_c \rightarrow P_h$
b. else $(P_k + P_l)/2 \rightarrow P_k$ for all k
6- If stop criterion is not satisfied then go to 2.

The stop criterion can be written as:

If $\sqrt{\sum((y_k - y_{critrion})^2/n)}$ is smaller than a pregiven parameter then algorithm converged and should be stopped. In which, y is the cost function value. Quality of reflection, extraction, contraction and shrinking can be understood through the algorithm above.

The simplex for two dimensional spaces (triangle) is displayed in Fig. 2. The projection new points are illustrated in this figure.

3. Proposed Augmented Downhill Simplex

As mentioned, the main idea of this modification on DSM is to hold all merits (speed and ease of the algorithm) besides improving the reliability. For this end, two changes in the algorithm are proposed here; first is that, logical iterative computation should be changed into vector computation; second is that, a random search operator should be added to the algorithm as a global search approach. Therefore, the proposed algorithm is shown here at a glance:

0- Initialize $0<\alpha<1$, $1<\gamma$ and $-1<\beta<0$
1- Initialize all vertexes of the simplex
2- Sort the simplex vertexes on depend on the cost function, and find P_h
3- Compute P_r , P_e , P_c by (1) and P_{rand} , as a random new position, and make a vector of them all
4- Sort a new points-vector on depend on the cost function and determine $P_{l\text{-}new}$ in the vector
5- $P_{l\text{-}new} \rightarrow P_h$ in the simplex
6- If the stop criterion (as it is described) is not satisfied then go to 2.

All other definitions are exactly similar to traditional DSM. Here, P_{rand} helps the algorithm to more explore the global space and perhaps would not be trapped in local minima.

To verify the effectiveness of this modification, several simulations have been done and the results are reported below.

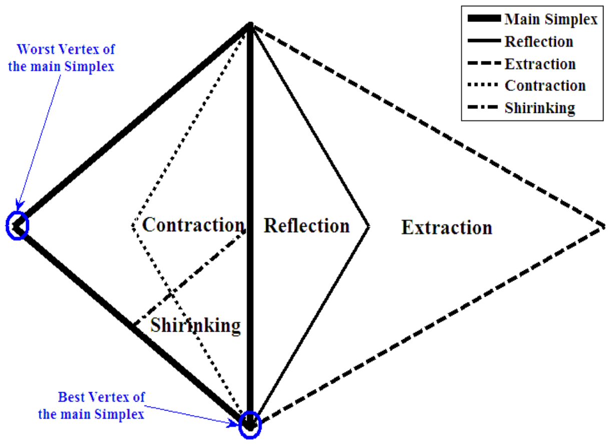

Fig. 2. The simplex and projection points

4. Simulation

Case 1: The presented method is applied to minimize an objective function called "biased peaks", (2).

$$\begin{aligned} z &= 3(1-x)^2 \exp(-x^2-(y+1)^2) \\ &-10(\frac{x}{5}-x^3-y^5)\exp(-x^2-y^2) \\ &-\frac{1}{3}\exp(-(x+1)^2-y^2)+6.552 \end{aligned} \quad (2)$$

In which, x and y are the variables of input space, and z is the output (cost value). To verify the efficiency and reliability of the proposed algorithm, the optimization procedures are executed thousand times and the average values are reported. Furthermore, traditional downhill simplex [1], random search [39] and steepest descent [40] algorithms are applied to minimize this objective function. For a graphical comparison, the final result of each run/method is shown via Fig. 3.a, (in the objective function contour). In which, black dots are the final results of minimizations by each method.

Diagram indicator of reliability is demonstrated through histogram (of thousand final results) in Fig. 3.b. The histogram can provide a proper estimation of each method's maneuver because the results are collected from 1000 runs of each method.

Figure 3, illustrates that each method has a specific 3ehaviour, i.e. steepest descent, tracks a similar path during minimization, random search makes a good diversity in all the space but cannot converge properly in the global minima, downhill simplex has better efforts but it is trapped in local minima, too. And finally the proposed algorithm benefits from the good properties of each method and makes more effective efforts. Thus, it managed to find the global minimum more than %68 of times; while other methods can only find this global minimum at most in %30 of times. This can be clarified through the histograms (Fig. 3.b). Although the proposed method still can be trapped in local minima, but its reliability is increased while its efficiency is reminded high.

To make a numerical comparison, the methods are compared via four important indices in table 1. As two indices of reliability the "*mean result*" that is the average value of all 1000 final results, and "*best result*" that is the best final value that methods can found are monitored. Also, as two indices of efficiency, the "*#Evaluated points*" that is the number of evaluating the cost function by methods and "*run time*" that is the average time of each run in milliseconds are observed.

table 1

A numerical comparison of the methods (Each number is an average value of 1000 runs, and 'run time' is in msec.)

Indices	Methods that applied to the optimization			
	Steepest Descent [40]	***Random Search [39]***	***Downhill Simplex [1]***	***Proposed ADSM***
Mean Result	4.2	4.3	3.7	1.3
Best Results	1e-3	1.2e-3	1e-3	9e-4
#Evaluated Points	20	40	23	31
Run Time	0.09	0.18	0.12	0.15

It can be seen in this table, that the proposed method, ADSM, can minimize the objective function better than other methods in similar situation. But, consider a situation in which we change the design parameters such that the number of evaluated point for all methods can would be equal; now, one question can be asked here; *"if the ADSM still works better?"*, as an answer to this question, it can be said that each algorithm determines the optimum number of needed cost-function-evaluation, freely. And this number is an average of 1000 times optimization process. If this value is restricted then the algorithms miscalculate the optimal point.

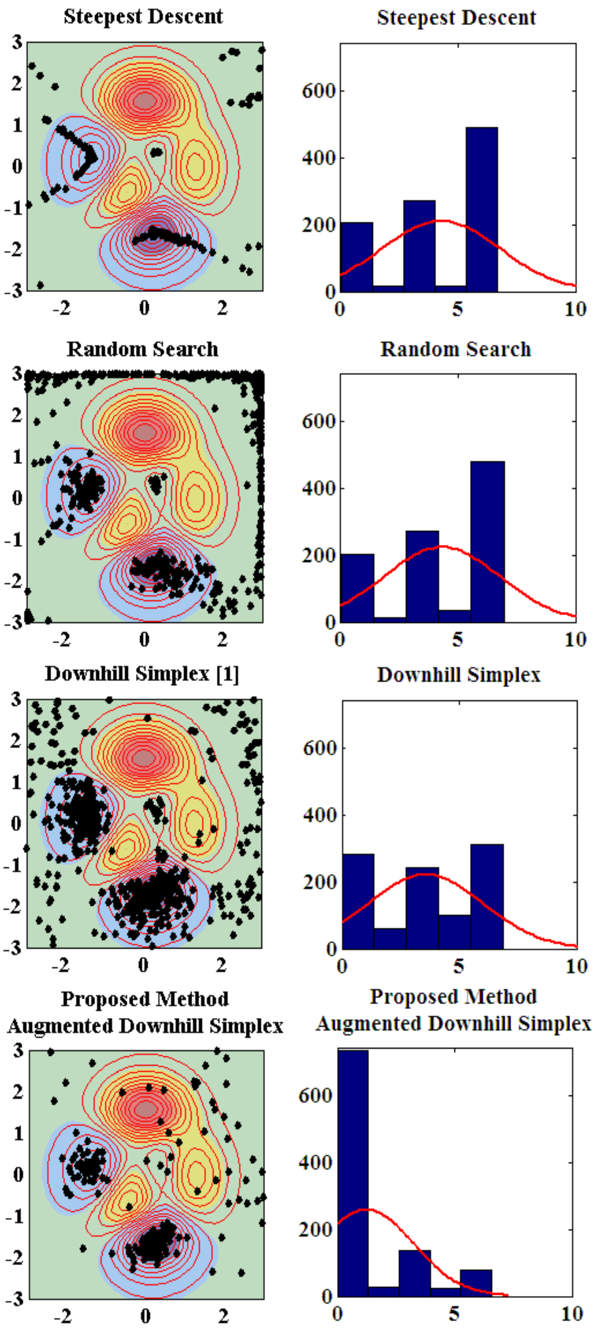

Fig. 3. A graphical comparison of four selected method (3.a) left side: the final results of all runs in contour space (3.b) right side: histogram of 1000 runs

Case 2: The presented method is also applied to minimize a non-smooth objective function, Fig.4.

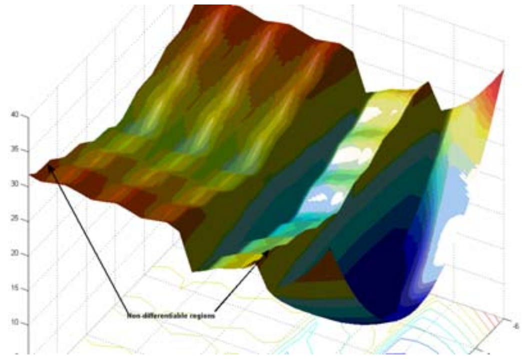

Fig. 4. Non-smooth objective function of case 2

Again, the optimization procedures are executed thousand times and the average values are reported. The final result of each run/method is shown in Fig. 5. This confirms that the reliability and efficiency of the proposed method is improved versus traditional method. As described before, the methods are compared via four important indices in table 2.

Because this cost function is not differentiable so steepest descent method would not be able to optimize this case.

Table 2

A numerical comparison of the methods in case 2.

Indices	Methods that applied to the optimization			
	Steepest Descent [40]	Random Search [39]	Downhill Simplex [1]	Propose d ADSM
Mean Result	-	**21.5**	**17.2**	**14.2**
Best Results	-	**13.0002**	**13.0001**	**13.0000**
#Evaluated Points	-	**40**	**25**	**33**
Run Time	-	**0.39**	**0.55**	**0.42**

In case 2, the superiority of the proposed method is concluded better. In fact, in the case of non-smooth, non-differentiable objective functions, differential-based method are deficient and would not be applicable. Also, derivative-free, local-search methods trap in local minima.

5. Conclusion

Here, a novel heuristic optimization method is introduced based on combination of downhill simplex and random search algorithms. The proposed Augmented Downhill Simplex Method (ADSM), is discussed in details and compare with three other commonly used algorithms; steepest descent, random search and traditional downhill simplex. Main idea of this combination is that, as people perhaps randomly take a look at sides while coming down from a mountain, in optimization simulations, a random operator is added to the proposed algorithm as a global search. To prove the improvement in proposed method, the ADSM is applied to optimization of two different objective functions. Simulation results confirm the merits of presented algorithm.

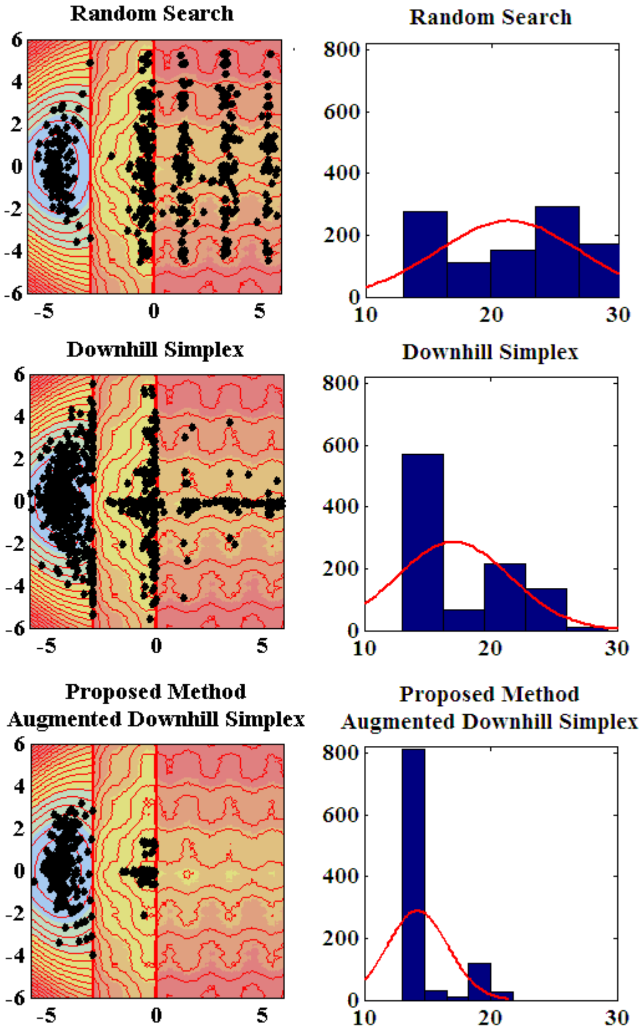

Fig. 5. A graphical comparison of selected method
(3.a) left side: the final results of all runs in contour space (3.b) right side: histogram of 1000 runs

References

[1] J.A. Nelder, R. Mead, “A simplex method for function minimization,” Comput. J., 7, 308- 313, 1965

[2] W. Spendley, G.R. Hext, F.R. Himswortf, “Sequential application of simplex designs in optimisation and evolutionary operation,”Technometrics 4, 441–462, 1962

[3] P.Wang, Q.Jin, S.Xu, H.Yang, H.F.Rashvand, “Efficient construction of irregular codes with midterm block length and near-Shannon performance,” IET Communications, 2010

[4] T.J. Wright, B.E. Parsons, J.A. Jackson, M. Haynes, P.J. Clarke, E.J. Fielding, P.C. England,"Source parameters of the 1 October 1995 Dinar (Turkey) earthquake from SAR interferometry and seismic bodywave modelling,"Earth and Planetary Science Letters, 172 (1999) 23–37

[5] H.Wang, L.Qian, E.Dougherty, “Inference of gene regulatory networks using S-system: a unified approach,” IET Systems Biology, 2008

[6] M. Musil, M. J. Wilmut, N. Ross Chapman, “A Hybrid Simplex Genetic Algorithm for Estimating Geoacoustic Parameters Using Matched-Field Inversion,” IEEE Journal of oceanic engineering, vol. 24, no. 3, July 1999

[7] M. Wachtler, A. Schweitzer, T. Gutmann, H. Breitzke, G. Buntkowsky,"Efficient analysis of 51V solid-state MAS NMR spectra using genetic algorithms,"Solid State Nuclear Magnetic Resonance, 35 (2009) 37–48

[8] N. Maehara, Y. Shimoda, "Application of the genetic algorithm and downhill simplex methods (NeldereMead methods) in the search for the optimum chiller configuration,"Thermal Engineering, 61 (2013) 433-442

[9] G.P. Kasprzyk, M. Jaskuła,"Application of the hybrid genetic-simplex algorithm for deconvolution of electrochemical responses in SDLSV method," Journal of Electroanalytical Chemistry, 567 (2004) 39–66

[10] M. Lehmann,"Determination and correction of the coherent wave aberration from a single off-axis electron hologram by means of a genetic algorithm,"Ultramicroscopy ,85 (2000) 165-182

[11] H. Sayoud, M.M. Hazem, K. Takahashi, "Hybridizing genetic algorithms for a survivability study of broadband communication networks: solution of the MSC," Communications, 2003. APCC 2003. The 9th Asia-Pacific Conference on , vol.3, no., pp.929,933, 21-24 Sept. 2003

[12] J. Yen, D. Randolph, J.C. Liao, B. Lee, "A hybrid approach to modeling metabolic systems using genetic algorithm and simplex method," Artificial Intelligence for Applications, 1995. Proceedings., 11th Conference on , vol., no., pp.277,283, 20-23 Feb 1995

[13] J. Yen, D. Randolph, B. Lee, J.C. Liao, "A hybrid approach to modeling metabolic systems using genetic algorithm and simplex method," Systems, Man and Cybernetics, 1995. Intelligent Systems for the 21st Century., IEEE International Conference on , vol.2, no., pp.1205,1210 vol.2, 22-25 Oct 1995

[14] X. Zhang, Y. Shen, S. Li, "A hybrid genetic algorithm for medical image registration with Downhill Simplex method," *Information, Computing and Telecommunication, 2009. YC-ICT '09. IEEE Youth Conference on* , vol., no., pp.291,294, 20-21 Sept. 2009

[15] Y.M. Jiang, Member, N. Ross Chapman, P. Gerstoft, “Estimation of Geoacoustic Properties of Marine Sediment Using a Hybrid Differential Evolution Inversion Method,” IEEE Journal of oceanic engineering, vol. 35, no. 1, January 2010

[16] V. C. Mariani, L. G. Justi Luvizotto, F. Alessandro Guerra, . L. dos Santos Coelho, "A hybrid shuffled complex evolution approach based on differential evolution for unconstrained optimization," Mathematics and Computation, 217 (2011) 5822–5829

[17] J. Kanka,"Design of turn-around-point long-period gratings in a photonic crystal fiber for refractometry,"Sensors and Actuators B, 182 (2013) 16– 24

[18] V. C. Mariani, L. S. Coelho,"A hybrid shuffled complex evolution approach with pattern search for unconstrained optimization, "Mathematics and Computers in Simulation, 81 (2011) 1901–1909

[19] F. M. Torres, E. Agichtein, L. Grinberg, G. Yu, R. Q. Topper,"A note on the application of the ‘‘Boltzmann simplex’’-simulated annealing algorithm to global optimizations of argon and water clusters,"Journal of Molecular Structure (Theochem), 419 (1997) 85–95

[20] E. Saboori, S. Parsazad, A. Sadeghi,"Improving the K-means Algorithm using Improved Downhill Simplex Search ," 2010 2nd International Conference on Software Technology and Engineering(ICSTE)

[21] A. Liesa, S. Lapinski, S. E. Dosso, “Bayesian Geoacoustic Inversion for the Inversion Techniques 2001 Workshop,”IEEE Journal of oceanic engineering, vol. 28, no. 3, July 2003

[22] S. Hartzell, P. Liu, “Calculation of earthquake rupture histories using a hybrid global search algorithm: application to the 1992 Landers, California, earthquake,” Physics of the Earth and Planetary Interiors 95 (1996) 79-99

[23] S. Kirkpatrick, C. D. Gelatt, M. P. Vecchi, “Optimization by Simulated Annealing,” Science, New Series, Vol. 220, No. 4598. (May 13, 1983), pp. 671-680

[24] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, Numerical Recipes in C, Second Edition, 1007pp. + xxvi (New York: Cambridge University Press, 1992).

[25] C. Aine, M. Huang, J. Stephen, R. Christner, "Multistart Algorithms for MEG Empirical Data Analysis Reliably Characterize Locations and Time Courses of Multiple Sources," NeuroImage 12, 159–172 (2000)

[26] M. Huang, C.J. Aine, S. Supek, E. Best, D. Ranken, E.R. Flynn,"Multi-start downhill simplex method for spatio-temporal

source localization in magnetoencephalography," Electroencephalography and clinical Neurophysiology, 108 (1998) 32–44

[27] Q. Duan, S. Sorooshian, Vijai Gupta,"Effective and efficient global optimization for conceptual rainfall-runoff models,"Water Resources Research,Volume 28, Issue 4, pages 1015–1031, April 1992

[28] T. Kanada, "Evaluation of spherical form errors-Computation of sphericity by means of minimum zone method and some examinations with using simulated data," Precision Engineering 17:281-289, 1995

[29] A. Krukowski, I. Kale,"Two Approaches for Fixed-point Filter Design: "Bit-Flipping" Algorithm and Constrained Downhill Simplex Method,"ISSPA'99, Brisbane, Australia, 22-25 August, 1999

[30] J. Qingwen, W. Pengjun, X. Shuzheng, Y. Huazhong, H.F. Rashvand, "Efficient construction of irregular LDPC codes with midterm block length and nearly optimal performance," Wireless Mobile and Computing (CCWMC 2009), IET International Communication Conference on , vol., no., pp.40,43, 7-9 Dec. 2009

[31] R. Bouzidi, Y. Lecieux, "A numerical method to optimize the design of a space inflatable membrane reflector," Acta Astronautica, 74 (2012) 69-78.

[32] B.C. Min, J. Lewis, E. T. Matson, A. H. Smith,"Heuristic optimization techniques for self-orientation of directional antennas in long-distance point-to-point broadband networks," Hoc Networks,11 (2013) 2252–2263

[33] P. Biernacki, S. Steinigeweg, A. Borchert, F. Uhlenhut, "Application of Anaerobic Digestion Model No. 1 for describing anaerobic digestion of grass, maize, green weed silage, and industrial glycerine,"Bioresource Technology, 127 (2013) 188–194

[34] N. Amoura, H. Kebir, S. Rechak, J.M. Roelandt, "Axisymmetric and two-dimensional crack identification using boundary elements and coupled quasi-random downhill simplex algorithms,"Engineering Analysis with Boundary Elements, 34 (2010) 611–618

[35] S. L. Garrison, B. J. Hardy, M. B. Gorbounov, D. A. Tamburello, C. Corgnale, B. A. vanHassel, D. A. Mosher, D. L. Anton, "Optimization of internal heat exchangers for hydrogen storage tanks utilizing metal hydrides,"International Journal of Hydrogen Energy, 37 (2012) 2850-2861

[36] S.E. Mansour, "Tuning Proportional-Integral controllers to approximate simplified predictive control performance," ISA Transactions, 48 (2009) 417-422

[37] P. J. Clarke, "Tectonic motions and earthquake deformation in Greece from GPS measurements," DPhil. University of Oxford**,** 1996, 274 pp

[38] C.K Chiang, S.H. Lai, "Fast Multi-Reference Frame Motion Estimation via Downhill Simplex Search," Multimedia and Expo, 2006 IEEE International Conference on , vol., no., pp.121,124, 9-12 July 2006

[39] R. Gran, "On the Convergence of Random Search Algorithm in Continuous Time with Applications to Adaptive Control," 1966

[40] C. A. Desoert, S. K. Mitrata, "Design Of Lossy Ladder Filters by Digital Computer," Ire Transactions on Circuit Theory, September 1961.