A Genetic programming for modeling Hadron-nucleus Interactions at 200 GeV/c

*Mahmoud Y. El-Bakry
University of Tabuk, Faculty of Science, Department of Physics, Tabuk, KSA.
** El-Sayed A. El-Dahshan, *** A. Radi
Ain Shams University, Faculty of Science, Department of Physics, Abbassia, Cairo, Egypt.
** Egyptian E-Learning University- 33 El-mesah St., El-Dokki- Giza- Postal code 12611.
*** The British University in Egypt (BUE).
M. Tantawy
*Ain Shams University, Faculty of Education, Department of Physics, Roxi, Cairo, Egypt.
*Moaaz A. Moussa
Buraydah Colleges, Al-Qassim, Buraydah, King Abdulazziz Road, East Qassim University, P.O.Box 31717, KSA.
moaaz2030@Yahoo.com

Abstract— Genetic programming (GP) is a soft computing search technique, which was used to develop a tree-structured program with the purpose of minimizing the fitness value of it. It is also a powerful and flexible evolutionary technique with some special features that are suitable for building a tree representation which is always the best solution for the problem we encounter. In this paper, GP has been used to describe a function that calculates charged and negative pions multiplicity distribution for Hadron-nucleus interactions at 200 GeV/c and also compared with the parton two fireball model (PTFM). GP calculations are in accordance with the available experimental data in comparison with the conventional ones (PTFM). Finally, the calculation results showed that the GP model is superior to the traditional techniques that we have ever seen so far.

Index Terms— Genetic programming (GP), machine learning (ML), pion production, multiplicity distribution.

IJSER

1 Introduction

High energies experimental data on hadron-nucleus (h-A) interactions are required for understanding high energy interactions. These data provide a useful link between hadron-hadron (h-h) interactions and the complex phenomena of nucleus-nucleus (A-A) interactions. These types of interactions investigate space time picture and highlight on phenomenon which doesn't exist in (h-h) such as gray particles, cascade, multi-collisions, etc. There are various models for (h-A) interaction like diffractive excitation model[1], collective tube model [2], quark model [3], energy flux cascade model [4], interanuclear cascade model [5], hydrodynamical model [6], multiple scattering model [7] and many others .

Conventional models like parton two fireball model (PTFM), treate nucleons as composite objects of loosely bound states of the spatially separated constituents (quarks) which in turn are composed of point-like particles (partons) [8]. This may allow one to consider the nucleons as consisting of a fixed number of partons. This nucleon structure has been used in different models [8-10] along with other assumptions to describe h-A interactions. PTFM, which is proposed by Hagedorn [11] has been used to explain the high energy interactions of hadrons and nuclei [12-18]. All these studies showed qualitative predictions of the measured parameters [19-23]. Extremely high energy collisions are required to get the fundamental particles close enough to study and understand the interactions between them [24–29].

Artificial intelligence techniques (or the machine learning) such as genetic programming (GP) are applicable for solving some problems in high energy physics [30–34]. The effort to understand the interactions of fundamental particles requires complex data analysis for which machine learning (ML) algorithms are vital. Machine learning (ML) algorithms are becoming useful as alternate approaches to conventional techniques [35]. The complex behavior of the h-A interactions due to the nonlinear relationship between the interaction parameters and the output often becomes complicated. In this sense, ML techniques such as artificial neural network [36], genetic algorithm [37], PYTHIA [38] and PHOJET [39] Monte Carlo models. The PHOJET model combines the ideas based on a dual parton model [40] on soft process of particle production and uses lowest-order perturbative QCD for hard process. PYTHIA on the other hand uses string fragmentation as a process of hadronization and tends to use the perturbative parton-parton scattering for low to high P_T particle production, and genetic programming [41] can be used as alternative tool for the simulation of these interactions [30–34, 42–47].

The motivation of using a GP approach is its ability to develop a model based entirely on prior data without the need of making underlying assumptions. Another motivation for applying such machine learning approach (e.g. GP) is simply the lack of knowledge (in most cases) about the mathematical dependence of the quantity of interest on the relevant measured variables [48].

In the present work, we illustrate the GP technique to model the multiplicity distribution of charged and negative pions for different beams at 200 GeV/c in hadron-nucleus collisions. The rest of the paper is organized as follows; Sec. 2 presents parton two fireball model PTFM at high energies for the multiparticle production in hadron-nucleus collisions Section 3 gives a review to the basics of the GP technique. Finally, the results and conclusion are provided.

2 Charged And Negative Pion Production In Hadron-Nucleus Collision Using PTFM

According to references [13, 15-18], the charged multiplicity distribution will be given by,

$$\mathrm{P}(\mathrm{n}_{\mathrm{ch}}) = \sum_{n=1}^{nch} \Phi(n)\Phi(n_{ch}-n) \tag{1}$$

$$; \quad n_{ch} = 2,4,6,........,Q/\varepsilon$$

Where, $\Phi(\mathrm{n}) = \sum_{n_0} \Psi(n_2)\,\mathrm{P}\,(n_o)$

, $\Psi(n_2)$ is the Poisson distribution of the form,

$$\Psi(\mathrm{n}_2) = \frac{\mathrm{N}^{\mathrm{n}_2}}{\mathrm{n}_2!}\, p^{\mathrm{n}_2}\ \mathrm{e}^{-\mathrm{NP}} \tag{2}$$

Where, N: is the number of pairs of created particles from one fireball ($N = \frac{n_0}{2}$), n_2 the number of pairs of charged pions, $n_2 = \frac{n-1}{2}$, P the probability that the pair of pions is charged.

The number of negative particles from one fireball equals the half of new created charged pions $n_- = \frac{n_{ch}}{2}$

The probability distribution of negative particles $\mathrm{P}(\mathrm{n}_-)$ is the same as the probability distribution of charged particles

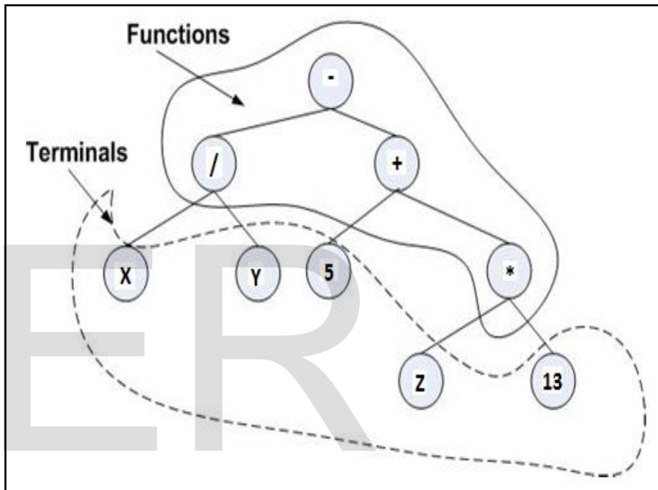
$$\mathrm{P}(\mathrm{n}_-) = \mathrm{P}(\mathrm{n}_{\mathrm{ch}} = 2\mathrm{n}_-) \tag{3}$$

$$; \quad n_- = 0,1,2,3,........,Q/2\varepsilon$$

3 Genetic Progrmming Overview

Genetic programming is an extension to Genetic Algorithms (GA). GA is an optimization and search technique based on the principles of genetics and natural selection. A GA allows a population composed of many individuals (chromosome) to evolve under specified selection rules to a state that maximizes the "fitness" (i.e. minimizes the cost function). The GP is similar to genetic algorithms but unlike the latter its solution is a computer program or an equation as against a set of numbers in the GA. A good explanation of various concepts related to GP can be found in Koza (1992) [41, 49].

In GP a random population of individuals (equations or computer programs) is created, the fitness of individuals is evaluated and then the 'parents' are selected out of these individuals. The parents are then made to yield 'offspring's' by following the process of reproduction, mutation and crossover. The creation of offspring's continues (in an iterative manner) until a specified number of offspring's in a generation are produced and further until another specified number of generations are created. The resulting offspring at the end of all this process is the solution of the problem. The GP thus transforms one population of individuals into another one in an iterative manner by following the natural genetic operations like reproduction, mutation and crossover. Each individual contributes with its own genetic information to the building of new ones (offspring) adapted to the environment with higher chances of surviving. This is the basis of genetic algorithms and programming. The representation of a solution for the problem provided by the GP algorithm is a tree (Fig. 1).



"Figure 1. Tree representation of the program $(X/Y)-(5+Z*13)$"

4 Results And Discussion

The GP is implemented using the experimental data to simulate multiplicity distribution of charged and negative pions for $p^{\pm}-Ar^{40}$, $p^{\pm}-Xe^{131}$, $p-Au^{197}$ and $p-He^{4}$ collisions at 200 GeV/c. The GP model was constructed with training sets and the accuracy was verified by the test sets. In order to generate the GP model we have implemented the GP steps (Fitness evaluation, reproduction, crossover and mutation) that were mentioned in Section 3. Table 1 lists the values of the control parameters and the set of function genes that are used in modeling the multiplicity distribution. The fitness function evaluates how accurate the mathematical model.

This discovered function has been used to predict the multiplicity distribution of pions for h-A interactions.

Simulation results based on GP model, for modeling the multiplicity distribution of charged pions for h-A interactions at 200 GeV/c (the training cases) are given in Fig. 2 (b, c, d,e) and Fig. 3 (b, c, d, e, f) for negative pions. While Figs. 2, 3 (a)

IJSER

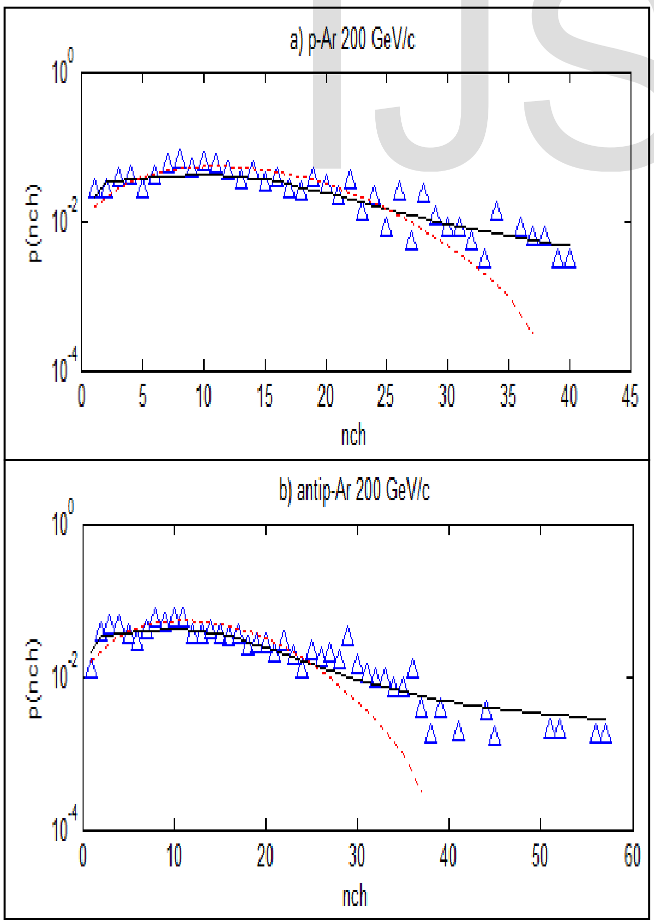
describe the predicted results for $p - Ar^{40}$ interaction at 200 GeV/c, we notice that the curves (for training and prediction cases) obtained by the trained GP model show a best fitting to the experimental data in all cases. Then, the GP model is able to exactly model the multiplicity distribution at 200 GeV/c for different beams in h-A collisions. If the large dataset is used in training, the best GP model is obtained.

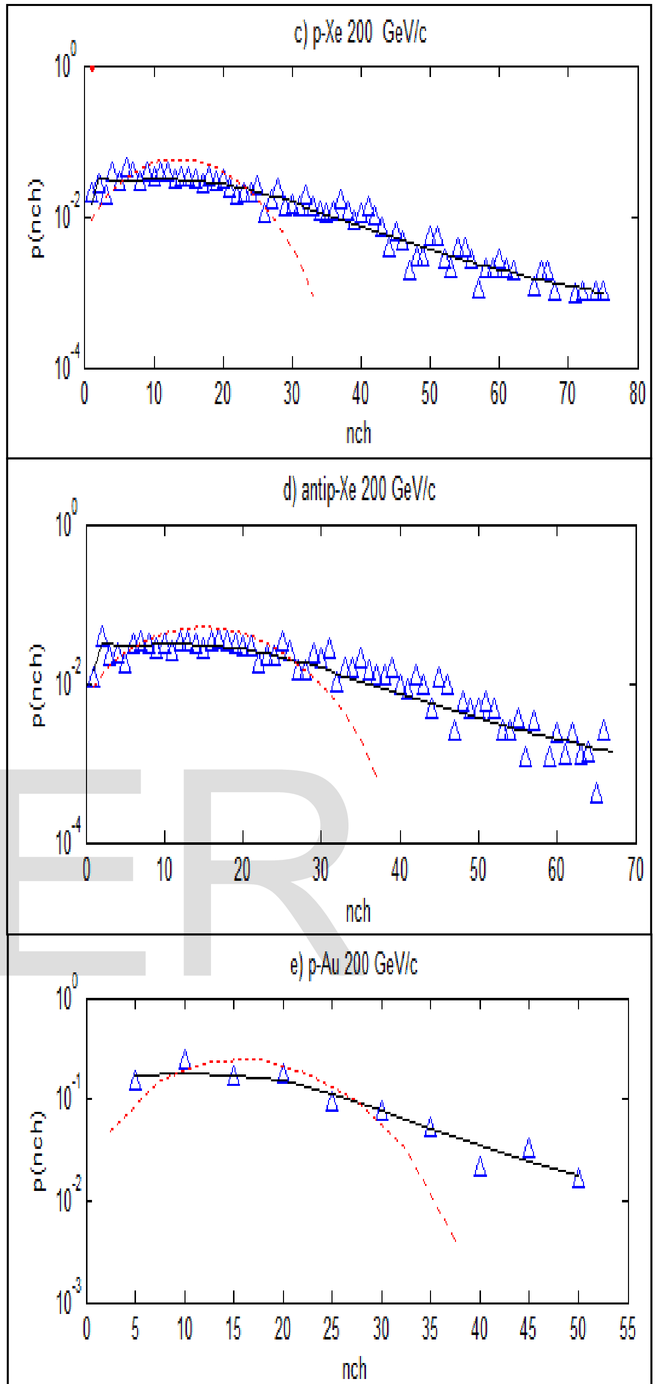
TABLE 1

VALUES OF THE CONTROL PARAMETERS USED IN MULTIPLICITY DISTRIBUTION

GP Parameters	Values
Generations	50
Populations	1500
Function set	*, /, -, +, log, power, sin, cos
Terminal Set	{constant, X,Y}
Fitness function	SSE
Selection method	Elites, rank and roulette
Mutation rate	0.01
Crossover rate	0.9
Fitness	0.9742

IJSER

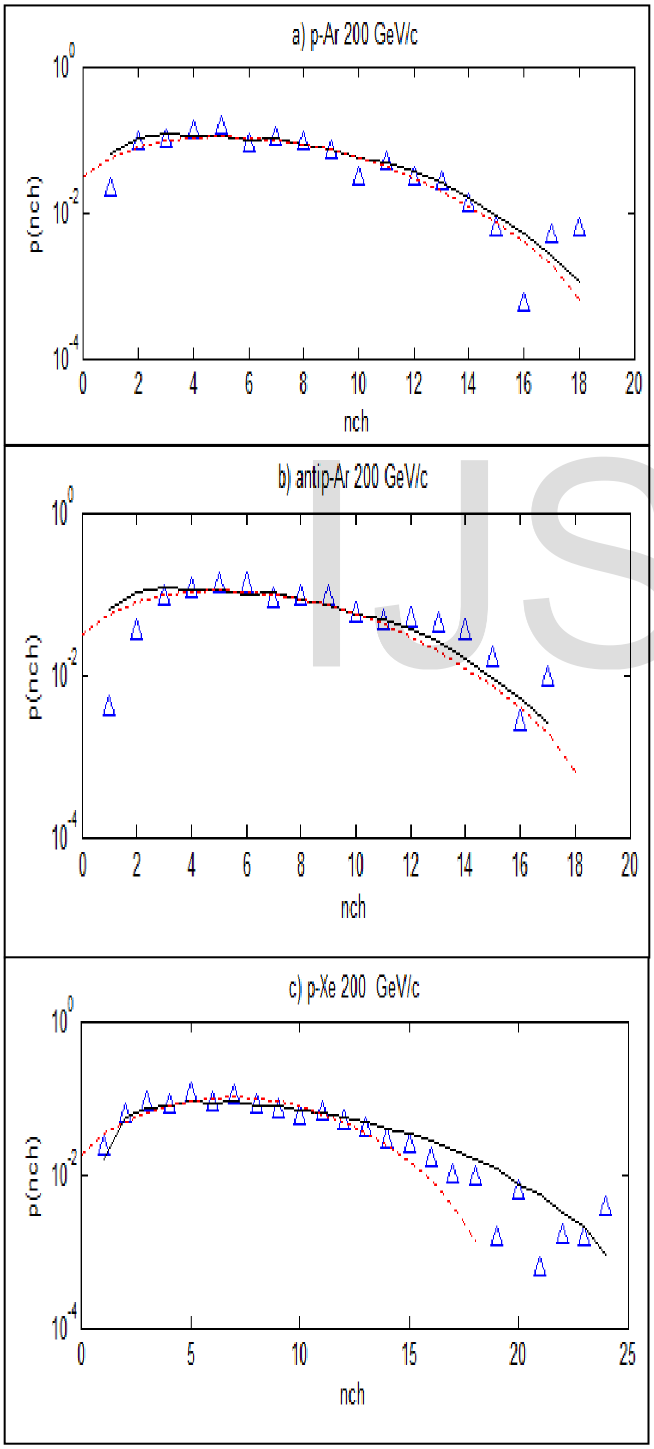


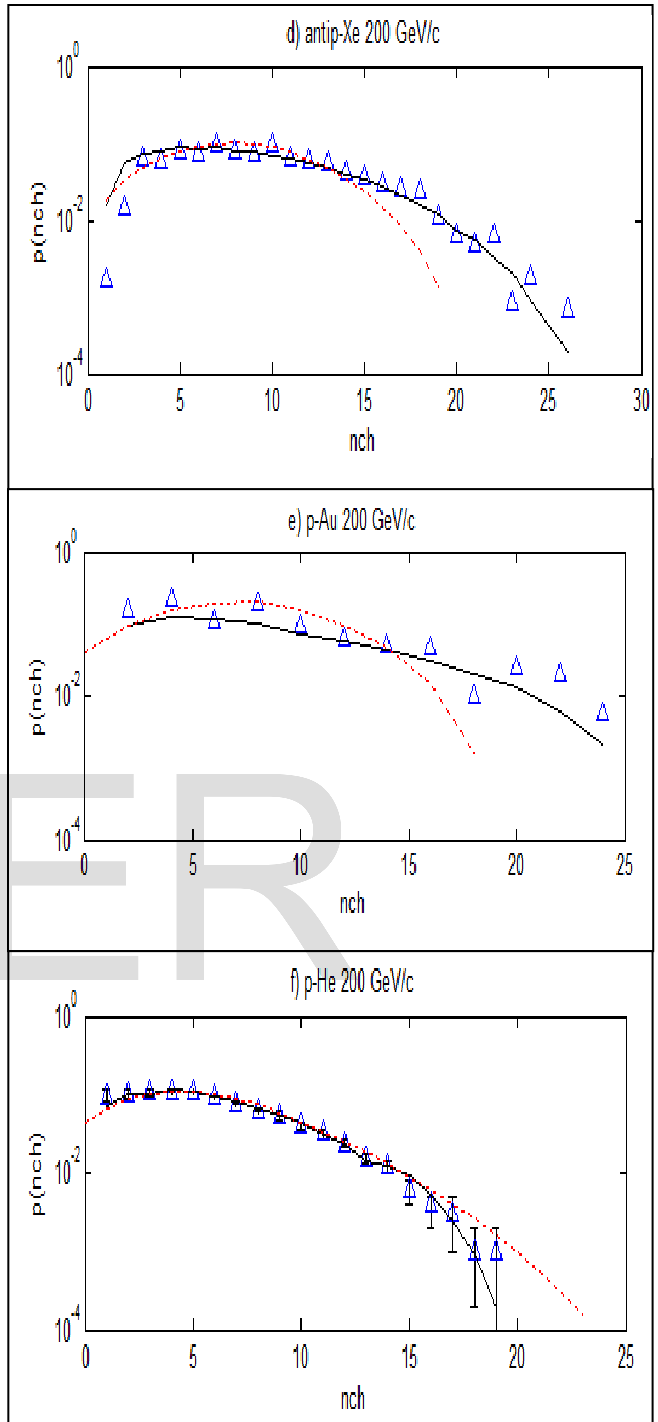


"Figure 2. Comparison between the experimental data, PTFM and simulated multiplicity distribution of charged pions $P(n_{ch})$ for h-A collisions at 200 GeV/c: (—) GP model, (……) PTFM, (△) experimental data"

5 CONCLUSION

Negative and charged pions multiplicity distributions, Eq. (1, 3), are calculated by PTFM for $p^{\pm} - Ar^{40}$, $p^{\pm} - Xe^{131}$, $p - Au^{197}$ and $p - He^{4}$ assuming ε is given by , $\varepsilon = a\,n_o + b$, Where, a = 0.04, b= 0.35 as in references [17, 18]. The results of these calculations are represented in figure 2 (a, b, c, d, e) and figure 3 (a, b, c, d, e, f) along with experimental data [50, 51] which show fair agreement with the corresponding experimental data. It can be seen from figs. (2, 3) that charged and negative pions multiplicity distributions are not in accordance with the experimental data for heavy nuclei although the situation becomes better for the light ones. The emission of secondary particles is assumed to follow a Poisson distribution. As mass number increases the multiplicity distribution is not broaden but its peak is shifted to high numbers.





“Figure 3. Experimental Data, PTFM, trained and predicted simulated multiplicity distribution of pions $P(n_-)$ for h-A collisions at 200 GeV/c: (—) GP model, (……) PTFM, (△) experimental data”

IJSER

Genetic programming, with its advantage of discovering mathematical equations (see APPENDIX), has been shown to be an efficient method for modeling the h-A interactions particularly above the pion production threshold. This paper presents an efficient approach for calculating the multiplicity distribution of charged and negative pions at 200 GeV/c through the obtained discovered functions as shown in appendix.

The discovered function shows an excellent match to the experimental data. Moreover, the discovered function is capable of predicting the experimental data that are not used in the training set. The present study has shown that the GP approach can be employed successfully to model the h-A interactions at high energies. Finally, we conclude that GP has become one of the important research areas in the field of hadron-nucleus collisions

APPENDIX

Our discovered function (for charged pion multiplicity distribution) is generated using the obtained control GP parameters as follows,

$$\mathrm{P(n_{ch})} = (((\mathrm{X}^3 + \mathrm{A})^{1.732021\mathrm{B}} * D^E)/F)/G$$

Where,

$$\mathrm{A} = (\mathrm{X}_1 - \mathrm{X}_3)/(10\mathrm{X}_3{}^{-1})^{10x_1^2}$$

$$\mathrm{B} = \cos(\mathrm{X}_1/(\log((\cos(\mathrm{X}_2$$

$$-(\log_2(\mathrm{X}_2) * (\mathrm{X}_3 + \mathrm{X}_2))))^b)))$$

$$\mathrm{b} = \log_2(\mathrm{X}_3) + 0.1\mathrm{X}_1$$

$$D = \cos((\sin(0.510701 - a^{bb}) + 10))$$

$$bb = ((\log_2(X_3) + 0.1\mathrm{X}_1)$$

$$a = \cos(\sin((0.510701 - (0.60512/(0.1\mathrm{X}_3)$$

$$* \cos(\log_2(0.29895/X_3) - 0.26714))) + 10))$$

$$E = ((\log_2(X_3) + 0.1\mathrm{X}_1)$$

$$F = \cos(\log_2(\log_2(\cos(\log_2(0.29895X_3{}^{-1}))))) + 1$$

$$G = \cos(\log_2(\log_2(\cos(\log_2(0.29895X_3{}^{-1}))))) + 1$$

for negative pion multiplicity distribution

$$\mathrm{P(n_{-})} = \sin((\mathrm{X}_1 - \cos(\mathrm{X}_3))/A)/(B)^{C+D}$$

Where,

$$\mathrm{A} = (10 - \sin(10\mathrm{g})) - (0.77965)^{\mathrm{X}_3}$$

$$\mathrm{g} = (e - \log_2(\mathrm{X}_2)^f) + \sin(\mathrm{X}_2)$$

$$\mathrm{e} = \cos(\cos(\mathrm{X}_1)) - (\mathrm{X}_3)^{\mathrm{X}_1}$$

$$\mathrm{f} = \log_2(\sin(\mathrm{X}_2)/\cos(\mathrm{X}_1))$$

$$\mathrm{B} = \log_2(\sin(\mathrm{X}_2)/\cos(\mathrm{X}_2))$$

$$\mathrm{C} = -(v/\mathrm{X}_3) + \mathrm{X}_1$$

$$v = B^{k+u}$$

$$u = t/\cos(\log_2(\mathrm{X}_2))$$

$$t = \sin(\sin(\cos(\log_2(\mathrm{X}_1 + s))))$$

$$s = B^r/\mathrm{X}_3$$

$$r = (n - o) + q$$

$$q = \sin(\cos(\mathrm{X}_3))/\cos(\log_2(\mathrm{X}_2)^p)$$

$$p = \sin(10^{\mathrm{X}_3} - \sin(\mathrm{X}_3))$$

$$o = \sin(\cos(\log_2(\mathrm{X}_1) * (\mathrm{X}_2 - \mathrm{X}_3))) + \mathrm{X}_1$$

$$n = \log_2(\mathrm{l})^m - \mathrm{X}_1$$

$$l = (2X_3 + 0.56464X_2 + 0.46078)/X_1$$

$$*(0.58503 - \mathrm{X}_3)$$

$$m = \sin(0.700228 - \sin(\mathrm{X}_2))$$

$$k = (\log_2(X_2)^j - X_1) - X_1 + (0.076463)$$

$$j = \sin(0.519666 - \sin(\sin(i)))$$

$$i = (\sin((X_2)^h))^h$$

$$h = \sin(\log_2(X_2) - \sin(X_3))$$

$$D = -0.83907/\cos(\log_2(X_2))$$

The actual parameters are, X_1 number of charged particles n_{ch}, n_-, X_2, lab momentum (P_L) and X_3, mass number (A).

REFERENCES

[1] A. Dar and J. Vary Phys Rev D 6, 2412 (1972).
[2] S. Fredriksson, Preprint Ref. TH. 2720-CERN, August (1979).
[3] Yoichiro Nambbu, The Conference of Quarks Scientific American (1976).
[4] K. Gottfried Phys. Rev. Lett 32, 957 (1974).
[5] V. S. Barashenkov et al., Sov. Phys. Usp 16, 31 (1973); Yu. P. Nikitin et al, Sov. Phys. Usp 20, 1 (1977).
[6] D. Amati et al., Nuovo Cimento. 26, 896 (1962); F. Zachariasen, Phys. Rep., C 2, 1 (1971).
[7] A. Capella and A. Krzywicki, Phys. Rev. D 18, 3357 (1978).
[8] R. P. Feynman, photon-Hadron Interactions (Benjamin, Reading, Massachusetts, 1972).
[9] J. Ranft, Phys. Lett, B 31, 529 (1970).
[10] E. Fermi, Prog. Theor. Phys. 5 , 570 (1950).
[11] R. Hagedorn, Nuovo Cimento Suppl. 3, 147 (1965); R. Hagedorn and J. Ranft, Nuovo Cimento Suppl.6, 169 (1968).
[12] M. Tantawy Ph. D. Dissertation, Rajasthan University (Jaipur-India) (1980).
[13] Ehab G. Abbas, M. Sc. Thesis Ain Shams University (Cairo-Egypt)

(2010)

[14] D.M.Habashy, Ph. D. Dissertation, Ain Shams University (Cairo-Egypt) (2011).

[15] S. Gamiel, M. Sc. Thesis Ain Shams University (Cairo-Egypt) (1997)

[16] Moaaz A. Moussa, M. Sc. Thesis Ain Shams University (Cairo-Egypt) (2009)

[17] M.Y. El-Bakry, M. Sc. Thesis, Ain Shams University (Cairo-Egypt) (1987).

[18] M. El-Mashad, Ph. D. Dissertation, Cairo University (Cairo-Egypt) (1994).

[19] M. Y. El-Bakry, Chaos, Solitons and Fractals 18, 995(2003).

[20] T. I. Haweel, M. Y. El-Bakry, K. A. El-Metwally, Chaos, Solitons and Fractals 18, 159 (2003).

[21] M. Tantawy, M. El-Mashad, S. Gamiel and M. S. El-Nagdy Chaos, Solitons and Fractals 13, 919 (2002).

[22] M. Tantawy, M. El Mashad, M. Y. Elbakry, The 3rd International Conference on Engineering Mathematics and Physics (ICEMP) Faculty of Eng., Cairo university, 23rd–25th Dec. (1997).

[23] M. Tantawy, M. El-Mashad, M. Y. El-Bakry. Indian J Phys.72A, 110 (1998).

[24] A. Augusto Alves et al. (LHCb Collaboration), J. Instrum. 3, S08005 (2008)

[25] G.L. Bayatian et al. (CMS Collaboration), *J. Phys. G: Nucl. Part. Phys.* 34, 995 (2007)

[26] [D.G. d'Enterria et al. (CMS Collaboration), *J. Phys. G: Nucl. Part. Phys.* 34, 2307 (2007)

[27] B. Alessandro et al. (ALICE Collaboration), *J. Phys. G: Nucl. Part. Phys.* 32, 1295 (2006).

[28] [F. Carminati et al. (ALICE Collaboration), *J. Phys. G: Nucl. Part. Phys.* 30, 1517 (2004)

[29] G. Aad et al. (ATLAS Collaboration), arXiv:0901.0512v4 [hep-ex]

[30] L. Teodorescu, D. Sherwood, Comput. Phys. Commun. 178, 409 (2008)

[31] L. Teodorescu, *IEEE T. Nucl. Sci.* 53, 2221 (2006)

[32] J.M. Link, *Nucl. Instrum. Meth. A* 551, 504 (2005)

[33] S. Yaseen El-Bakry, Amr Radi, *Int. J. Mod. Phys. C* 18, 351 (2007)

[34] E. El-dahshan, A. Radi, M.Y. El-Bakry, *Int. J. Mod. Phys. C* 20, 1817 (2009)

[35] S. Whiteson, D. Whiteson, *Eng. Appl. Artif. Intel.* 22, 1203 (2009)

[36] S. Haykin, Neural Networks: A Comprehensive Foundation (*IEEE* Press and Macmillan College Publishing Company, New York, NY, 1994)

[37] J.H. Holland, Adaptation in Natural and Artificial Systems (University of Michigan Press, Ann Arbor,1975)

[38] T. Sjostrand, et al., Computer Physics Commun. 2001 135 238; T. Sjostrand and M. van Zijl, Phys.Rev. D 36 (1987) 2019; T. Sjostrand and P. Skands, Eur. Phys. J. 2005C 39 129; T. Sjostrand and P. Skands, 2006 JHEP 0605 026.

[39] R. Engel, Z. Phys. 1995C 66 203; R. Engel, J. Ranft and S. Roesler, 1995 Phys. Rev. D 52 1459.

[40] A. Capella, et al., Phys. Rep. 1994 236 225.

[41] J.R. Koza, Genetic Programming: On the Programming of Computers by means of Natural Selection (The MIT Press, Cambridge, MA, 1992)

[42] K. Cranmer, R.S. Bowman, Comput. Phys. Commun.167, 165 (2005)

[43] M.Y. El-Bakry, A. Radi, *Int. J. Mod. Phys. C* 18, 329 (2007)

[44] M.Y. El-Bakry, K.A. El-Metwally, *Chaos Soliton. Fract.* 16, 279 (2003)

[45] K.A. El-Metwally, T.I. Haweel, M.Y. El-Bakry, *Int. J. Mod. Phys. C* 11, 619 (2000)

[46] E. El-dahshan, A. Radi, M.Y. El-Bakry, *Int. J. Mod. Phys. C* 19, 1787 (2008)

[47] A.E. Eiben, J.E. Smith, Introduction to Evolutionary Algorithms (Springer, Berlin, 2003)

[48] H. Etemadi, A.A.A. Rostamy, H.F. Dehkordi, *Expert Syst.* Appl. 36, 3199 (2009)

[49] M. Wolter, *Phys. Part. Nucl.* 38, 255 (2007)

[50] D. H. Brick et al., Phys. Rev. D 39, 2484 (1989).

[51] W. Bell et al., Phys. Lett. B128, 349 (1983).

IJSER