Evolutionary algorithm with α-stable mutation

Andrzej Obuchowicz
Institute of Control and Computation Engineering
University of Zielona Góra
Podgrna 50, 65-246 Zielona Góra, Poland
e-mail: A.Obuchowicz@issi.uz.zgora.pl

Przemysław Prętki
Institute of Control and Computation Engineering
University of Zielona Góra
Podgrna 50, 65-246 Zielona Góra, Poland
e-mail: P.Pretki@issi.uz.zgora.pl

Abstract— **One-dimensional Symmetric α-Stable ($S\alpha S$) mutations are applied to phenotypic evolutionary algorithms. The exploration and exploitation abilities of $ESSS_\alpha$ algorithm, using the proposed mutations, are analyzed by a set of simulation experiments. The obtained results, showed that Normal or Cauchy mutation may be challenged.**

I. Introduction

Most applications of Evolutionary Algorithms (EAs) which use the floating point representation of population individuals use the Gaussian mutation as a mutation operator [1], [2], [4], [5], [6], [11], [18]. A new individual x is obtained by adding a normally distributed random value to each entry of a selected parent y:

$$x_i = y_i + N(0, \sigma_i), \quad i = 1, \ldots, n. \tag{1}$$

The choice is usually justified by the Central Limit Theorem. Mutations in nature are caused by a variety of physical and chemical factors that are not identifiable or measurable. These factors are considered as independent and identically distributed (i.i.d.) random perturbations. The Generalized Central Limit Theorem states that the only possible nontrivial limit of normalized sums of i.i.d. terms is Lévy-stable [10], called also α-stable or just stable in the mathematical literature [3], [12], [22], [25]. If the Lindeberg condition is obeyed, i.e., the first two absolute moments are finite, then the Lévy-Stable Distribution (LSD) reduces to the Gaussian distribution. The lack of closed form formulas for probability density functions (pdfs) for all but three LSDs (Gaussian (GD), Cauchy (CD) and Lévy (LD) distributions) has been a major drawback in the use of LSDs by practitioners. Fortunately, there exist algorithmic formulas for simulating Lévy-stable variables [24] as well as computer programs to compute Lévy-stable densities, distribution functions and quantiles [12].

The suggestion that the application of LSDs other than the GD and CD can be very attractive for evolutionary algorithms with the floating-point representation of individuals was first introduced by Gutowski [8], but this idea has not be pursued so far. The aim of this work is to investigate the effectiveness of a simple model of asexual phenotypic evolution with mutations based on LSDs.

This work is organized as follows. First, LSDs are defined in Section II and their some properties are described. The main part containing the experimental studies is presented in Sections IV, V, where local and global convergence of $ESSS_\alpha$ algorithm are considered. Finally, the work is concluded.

II. Stable distribution

Definition 1: A random variable X is *stable* or *stable in the broad sense* if for X_1 and X_2 independent copies of X and any positive constants a and b,

$$aX_1 + bX_2 = cX + d \tag{2}$$

for some positive c and some $d \in \mathbb{R}$.
The random variable is strictly stable or stable in the narrow sense if (1) holds with $d = 0$ for all choices of a and b.
Due to the lack of closed form formulas for densities, the LSD can be most conveniently described by its characteristic function (ch.f.) $\phi(k)$ — the inverse Fourier transform of the pdf. The ch.f. of the LSD is parameterized by a quadruple $(\alpha, \beta, \sigma, \mu)$ [24], where α ($0 < \alpha \leq 2$) is a stability index (tail index, tail exponent or characteristic exponent), β ($-1 \leq \beta \leq 1$) is a skewness parameter, σ ($\sigma > 0$) is a scale parameter and μ is a location parameter. There are a variety of formulas of the LSD ch.f. in the relevant literature. This fact is caused by a combination of historical evolution and numerous problems that have been analyzed using specialized forms of LSDs. The most popular formula of the ch.f. of $X \sim S_\alpha(\beta, \sigma, \mu)$, i.e., a Lévy-stable random variable with parameters α,β,σ and μ, is given by [23]:

$$\phi(k) = \exp\left(-\sigma^\alpha |k|^\alpha \left\{1 - i\beta \mathrm{sign}(k) \tan\left(-\frac{\pi\alpha}{2}\right)\right\} + i\mu k\right) \tag{3}$$

when $\alpha \neq 1$, and

$$\phi(k) = \exp\left(-\sigma|k| \left\{1 + i\beta \mathrm{sign}(k) \frac{2}{\pi} \log|k|\right\} + i\mu k\right) \tag{4}$$

when $\alpha = 1$.

A. Selected properties of the LSD

The classical Central Limit Theorem says that the normalized sum of i.i.d. with finite variance converges to a normal distribution. The Generalized Central Limit Theorem shows that if the finite variance assumption is dropped, the only possible resulting limits are stable.

Theorem 1: **Generalized Central Limit Theorem** [10] Let $X_1, X_2, X_3, \ldots$ be an i.i.d. sequence of random variables.

There exist constants $a_n > 0, b_n \in \mathbb{R}$ and a nondegenerate random variable Z with

$$a_n(X_1 + X_2 + \ldots + X_n) - b_n \longrightarrow Z$$

if and only if Z is Lévy-stable for some $0 < \alpha \leq 2$.
In this way, a basic property of stable laws is that sums of Lévy-stable random variables are Lévy-stable.
While there are no explicit formulas for general stable densities, a lot is know about their theoretical properties. The most basic fact is the following:

Theorem 2: All (non-degenerate) stable distribution are continuous distribution with infinitely differentiable densities.
To distinguish between the densities and cumulative distribution functions, $f_\alpha(x|\beta, \sigma, \mu)$ will denote pdf and $F_\alpha(x|\beta, \sigma, \mu)$ will denote the cdf of $S_\alpha(x|\beta, \sigma, \mu)$ distribution.
When $\alpha = 2$, the normal distribution has well understood asymptotic tail properties. Te tail probabilities in the non-Gaussian cases are know asymptotically.

Theorem 3: **Tail approximation** [12]. Let $X \sim S_\alpha(\beta, \sigma, \mu)$ with $0 < \alpha < 2, -1 < \beta < 1$. Then as $x \longrightarrow \infty$,

$$P(X > x) \sim \quad \sigma^\alpha c_\alpha (1 + \beta) x^{-\alpha} \tag{5}$$

$$f_\alpha(x|\beta, \sigma, 0) \sim \quad -\alpha\sigma^\alpha c_\alpha (1 + \beta) x^{-(\alpha+1)} \tag{6}$$

where $c_\alpha = sin(\pi\alpha/2)\Gamma(\alpha)/\pi$
A general distribution is said to be heavy tailed if its tails are heavier than exponential. For $\alpha < 2$, stable distribution have one tail (when $\alpha < 1$ and $\beta = \pm 1$) or both tails (all other cases) that are asymptotically power laws with heavy tails. A consequence of heavy tails of LSDs is that not all moments exist. In most statistical problems, the first moment $E(X)$ and variance $\mathrm{Var}(X) = E(X^2) - (E(X))^2$ are routinely used to describe a distribution. In the case of LSDs such a representation is not useful since we have

$$E(X^p) = \int_{-\infty}^{\infty} x^p f(x) dx = \mu < +\infty \Leftrightarrow 0 < p < \alpha. \tag{7}$$

Thus, the second moment exists only for $\alpha = 2$, the first moment exists for $1 < \alpha \leq 2$ and is equal to the location parameter μ [12].

Usually, the used random vector in mutate operator must fulfill some basic requirements. It is reasonable to postulate that - at least initially - no preference of a certain direction should be given. This request leads to the property that probability density function should be symmetric with respect to the origin. For this reason, we consider only spherically α-symmetric distribution $S\alpha S(\sigma)$ e.i.:

$$S\alpha S(\sigma) = S\alpha(0, \sigma, 0) \tag{8}$$

In this case, characteristic function is reduced to following form:

$$\phi(k) = \exp(-\sigma^\alpha |k|^\alpha) \tag{9}$$

III. Evolutionary Algorithm with Soft Selection and α-stable mutation - $ESSS_\alpha$

The $ESSS_\alpha$ algorithm is based on the simple model of phenotypic asexual evolution [7]. The model describes the evolution of η-individual population with nonoverlapping generations. Each individual is characterized by an n-dimensional vector of traits and a fitness value assigned to the vector. The searching process consists in generating a sequence of η-element populations. A new population $G(t+1)$ is created based only on the previous population $G(t)$. In order to generate a new element x_k^{t+1}, a *parent* element is selected and mutated. Both selection an mutation are random processes. Each element x_k^t can be chosen as a parent with the probability proportional to its fitness (*the roulette method*). A new element x_k^{t+1} is obtained by adding a α-stable distributed random value to each entry of the selected parent:

$$(x_k^{t+1})_i = (x_k^t)_i + S\alpha S(\sigma) \quad i = 1, \ldots, n. \tag{10}$$

An optimization algorithm corresponding to this model is used in a real-valued unconstrained optimization. The original ESSS has appeared to be an efficient method of saddle crossing in both one- and high-dimensional search space [16].
Due to the interrelated stochastic processes, the convergence analysis of the algorithm, is an extremely complicated task, if not impossible. Therefore, some experimental results obtained for simple evolutionary models, may prove to be very valuable. They may support us with the indispensable knowledge about the processes occured during execution an optimization task. In the following chapters authors attention is focused on the factors, which have a direct influence on local convergent of $ESSS_\alpha$ algorithm.

IV. Local convergence

The complexity of the local convergence analysis derives from the fact that we have to deal with problems for which simple mathematical models do not exist. If one wants to evaluate the efficiency of evolutionary algorithm in local optimization tasks, taking two of the most obvious criteria into consideration seems to be essential:

- the convergence velocity
- the quality of the finding solutions.

The velocity of local convergence of simple (1+1)ES and (1,λ)ES, was considered by Rudolph [20]. He proved that the order of local convergence is identical for Gaussian and spherical Cauchy distribution, whereas non-spherical Cauchy mutation leads to slower local convergence.
Very interesting results were obtained by Iwona Karcz-Dulęba [9], who before anyone else, made an attempt at convergence analysis of ESSS algorithm. She proposes to consider the evolution dynamic in the state space of the population. In spite of using very simple evolutionary model, she proved that, in local optimization tasks, population of individuals is convergent not to the local optimum, but to the point located in some distant from it.

When ones want to enumerate all factors which have a significant influence on the local convergence of evolutionary algorithms, one cannot help notice results obtained by Obuchowicz [15]. He showed that the mutation operator combined with multidimensional search space is responsible for creation so-called surrounding effect. Due to the fact that the surrounding effect is crucial to the quality of finding solutions, authors decided to look at this problem more carefully in the next subsection.

A. Surrounding effect

The problem of surrounding effect is related to the probability that the distance from the mutated point x and its offspring y will be in the range $\|x-y\| \in [r, r+dr]$. Although the pdfs of α-stable distributions have their maxima at mutated points, it is easy to prove (c.f. [13], [14]) that the most probable distance of the offspring is near zero only in the case of a one-dimensional mutation. In the case of an n-dimensional mutation, the most probable distance increases with n, in the case of a Gaussian mutation it is proportional to the norm of the standard deviation vector and to $\sqrt{n-1}$. This fact, called the *surrounding effect* [14], formed the basis for analytical investigations of Evolutionary Strategies (ES) [19], [21], and a simulation analysis of the ESSS and EP algorithms with Gaussian and Cauchy mutations [15], [16]. The surrounding effect decreases the sensitivity of an EA to narrow peaks of the fitness function while increasing the dimension of the search space. In order to show surrounding effect a simple experiment has been done. Figure 1 shows histograms of the distances between the origin and 10^6 points mutated with chosen distributions SαS for some space dimensions n, and for different indexes α. Histograms 1 very clearly show that the region of the most probably points location, depends on the dimension of search space n as well as stability index α. Looking at the histograms related to $\alpha = 0.5$ we must to be aware of a situation in which only wide spread areas of a search space are sampled, what stands for chaotic behavior of population.

B. The upper limit of local convergence - Dead zone

In the aftermath of the surrounding effect, arises an area, where probability of finding an individual, is close to zero. It means that population in its stable state fluctuates on the border of this region, and only on the very rare occasion explores its interior. In the case of one dimensional search space, the surrounding effect does not exist. Nevertheless, one can notice that population is not convergent to the local optimum, which similarly as previously is surrounded by *dead zone*. To gain an insight into this phenomenons, one can be aware that the population is effected by two forces, incessantly. The first one derives from selection operator, which is responsible for choice the direction of searching process. The second one is determined by mutate operator, which is responsible for exploitation mechanism.

Far from the local optimum, this forces complement one another, and move population in a direction of the better fitted

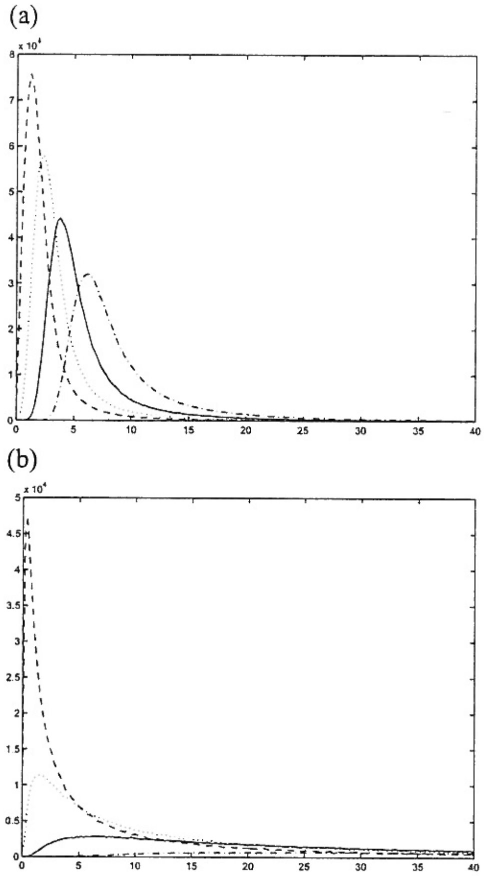

Fig. 1. Histograms of the distances between the base point and 10^6 points mutated according to (a) $S_{1.5}(1)$ and (b) $S_{0.5}(1)$ mutations; $n = 2$-dashed line, $n = 4$-doted line, $n = 8$-solid line, $n = 16$-dash-dotted line

areas. Unfortunately, as distant from local optimum decreases, the mutate operator restrains population from further progress. To better outline this situation, the upper limit of probability of successful mutations has been estimated. The successful mutation term relates to the situation, in which the best fitted individuals has been improved by mutating base population. In the case of two well know stable distributions ($\alpha = 2, \alpha = 1$) such probabilities can be expressed as:

$$P(f(x'_B) < f(x_B)) < 1 - \left[1 - 0.5\mathrm{erf}\left(\frac{x_B\sqrt{(2)}}{\sigma}\right)\right]^{\eta} \quad (11)$$

for Gaussian and

$$P(f(x'_B) < f(x_B)) < 1 - \left[1 - \frac{1}{\pi}\arctan\left(\frac{2x_B}{\sigma}\right)\right]^{\eta} \quad (12)$$

Cauchy mutation, where x_B and x'_B mean the best elements in the base and temporary population respectively, η stands for population size.

Above formulas concern one-dimensional search space, and remain valid only if the following condition is obeyed:

$$|x_1| < |x_2| \Rightarrow f(|x_1|) < f(|x_2|) \quad (13)$$

Naturally, as closer population is situated to the local solution as probability of successful mutation decreases. If one look

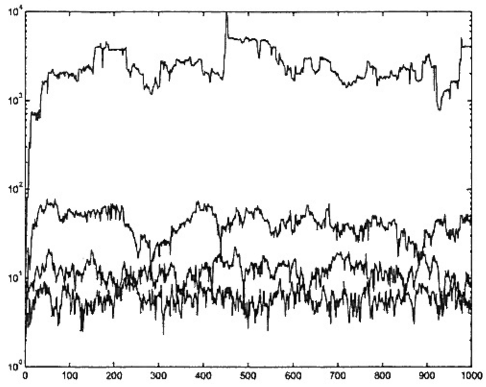

Fig. 2. ESSS process for $f_{sph}(X)$ stared in the optimum point. The distance between the best point in the population and the optimum point vs. iteration. Stable index $\alpha = 0.5, 1, 1.5, 2$ from the top to the bottom curves respectively

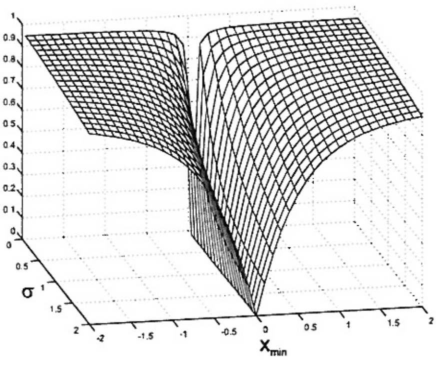

Fig. 3. The probability of successful mutation for $ESSS_1$ algorithm vs. value of scale parameter σ and distant of the best element in the population X_{min} from the local optimum $X = 0$.

at the figure 3, it becomes obvious, from where mentioned *dead zone* comes. When the population is located suitably near to the local optimum, to generate better solutions is almost impossible. This means that the population is pushed back beyond area, where the probability of successful mutation is quite slight. In this way the population starts to fluctuate around local optimum. To gain an insight into problem of dead zone, following experiment has been done. The population of 20 individuals has been considered in the one dimensional search space with simple subject $f(x) = x^2$. As a start point of the $ESSS_\alpha$ algorithm, the global optimum point $x = 0$ has been chosen. The smallest distance from global optimum point during 1000 epochs has been presented on the figure 2. The experiment has been executed for different stable mutations. On the basis of above experiment, it becomes obvious that quality of the finding solution is very tightly connected to the distribution of random vector in mutate operator. Generally, as smaller stable index α is, as population managed by $ESSS_\alpha$ algorithm, places its individuals further from global optimum. Since, as stable distribution is defined also by scale parametr, the next chapter has been devoted to examine influence of both parameters on quality of $ESSS_\alpha$'s local convergence.

C. Quality of the finding solution

In order to exemplify the influence of a mutation step σ, and the index of stability α on the quality of finding solutions, let us look at the following experiment, which may prove helpful. Considering the one dimensional sphere function $f(x) = x^2$, the parameters are as follows: $\eta = 20$, $t_{max} = 10000$, $x_0 = 0$. Figures 4 illustrate how far from the optimum point the population fluctuates in its stable state for $\alpha = 0.5, 0.6, ..., 1.9, 2$ and $\sigma = 0.001, 0.002, ..., 0.009, 0.01, 0.02, ..., 0.09, 0.1$. Accuracy

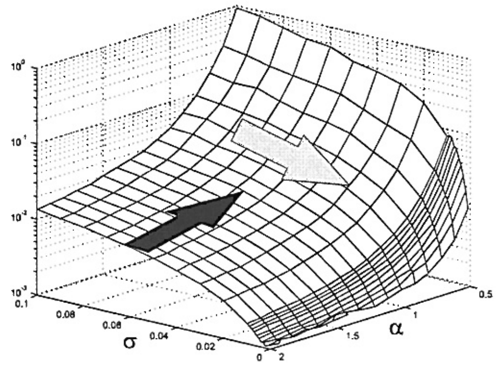

Fig. 4. The size of the *dead zone*, estimated on the basis of experiment described in the text, for various parameters of stable mutation. The black and gray arrows indicate the direction, where better global and local $ESSS_\alpha$'s convergence can be achieved respectively

of the finding solution is sensitive to index α as well as scale parameter σ. The best results (the smallest *dead zone*) can be obtained for $\alpha = 2$. As stable index α decreases as the population places its element farther from the optimum point. On the basis of the experiment, one can notice that fixed local convergence accuracy can be obtained in two different manners. Firstly for any stable index α can be chosen some scale parameter σ such as, $ESSS_\alpha$ algorithm will be endowed with fixed local convergence accuracy. One the other hand, for any scale σ may be selected particular index α, such as considered condition will be fulfil as well.

V. Global convergence

In order to analyze influence of the α-stable mutation on the global convergence, the simple saddle crossing problem was considered. The objective function is composed of two Gaussian peaks. The lowest one posses its optimum at the point $x = 0$, and the global optimum is located at the point $x = 1$. The goal is to cross the saddle between both peaks. We assume that it is done, when one of the individuals gains its fitness value greater than 0.6. The initial point of searching for the tested algorithm is chosen in the local optimum of the function $\Phi_{sc}(x)$. Other algorithm parameters were chosen as follows: the population size $\eta = 20$, the maximum number of epoch $t_{max} = 10^6$, stable index $\alpha = 0.5, 0.6, ..., 1.9, 2$, and scale parameter $\sigma = 0.001, 0.002, ..., 0.009, 0.01, 0.02, ..., 0.09, 0.1$.

Observation 1: On the figure 5, one can say, that for indexes α smaller than 1.2, the effectiveness of saddle crossing

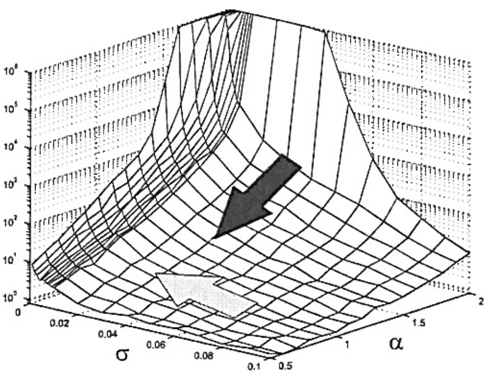

Fig. 5. The mean number of epochs needed to cross the saddle for various parameters of stable mutation. The black and gray arrows indicate the direction of better global, and local $ESSS_\alpha$'s convergence respectively.

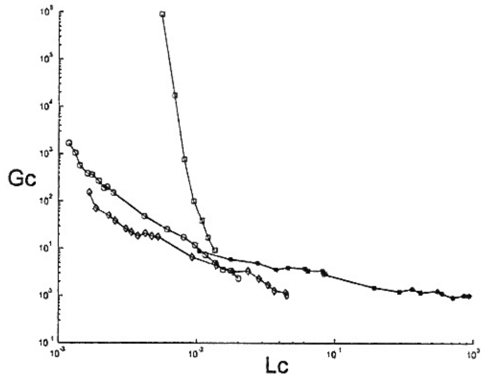

Fig. 6. The mean number of epochs needed to cross the saddle G_C and the precision of finding solution G_L for various indexes α, and scale parameters σ. ($\alpha = 0.5$ stars, $\alpha = 1$ diamonds, $\alpha = 1.5$ circles, $\alpha = 2$ squares)

does not significantly depend on the scale parameter σ. On the other hand, for bigger α relationship between the velocity of saddle crossing and scale σ is much stronger. In fact, for big stable indexes α, even small changes of scale parametr σ, may causes loss of global convergence ability.

The most interesting observation relating to described in this section results, has been stressed with the help of two arrows attached to the figure 5. If one looks at the figure 5 more precisely, it becomes obvious, that for each index $alpha$, exists a *path of better global convergence*. Generally it can be noticed that when α decreases, $ESSS_\alpha$ algorithm gains better global convergent ability. But based on the results of first experiment we know, that for smaller indexes α, $ESSS_\alpha$ algorithm has a little changes to locate optimum point precisely. On the other hand we also know that the local convergent can be improved by decreasing scale parameter σ (the direction pointed out by gray arrow on the figure 5).

VI. Effectiveness of $ESSS_\alpha$ algorithm in global optimization tasks

The results obtained for two previous experiments could be juxtapose with each other on the figure 6, where results for the most interesting $ESSS_\alpha$ cases has been presented. Results of the first and second experiment have been put on vertical and horizontal axis respectively. In this sense, as closer point is located to the vertical axes , as better local convergent has been obtained by $ESSS_\alpha$ algorithm. Thereby, points situated closer to the horizontal axis, correspond to the parameters of $ESSS_\alpha$ algorithm with better global convergent abilities. Points lied on the same line correspond to various scale parameters of certain α-stable mutation. The first point on the right side corresponds to the value $\sigma = 0.1$, the second one $\sigma = 0.09$ and so on. The best algorithm of global optimization ought to be characterized by the shortest time needed to escape from local optimum and the highest accuracy in locating global solution, simultaneously. For that reason one can assume that parameters of stable mutation corresponding to the closest point to the origin, define mutation with the best optimization abilities. Taking into consideration above remarks, one can notice that for considered problems, the best global optimization abilities possess Cauchy distribution. $ESSS_1$ corresponding to evolutionary algorithm with Cauchy mutation posses slight worst local convergence only for the upper range of scale parameters. For $\sigma < 1.3$ Of course this kind of results very strongly depend on the problem formulation.

VII. Multivariate α-stable mutation

If $\mathbf{X} = (X_j \sim S_\alpha S | j = 1, 2, ..., n) \sim \mathbf{S_\alpha S}$ is a sample from a stable law, its characteristic function is given by

$$\phi(\mathbf{k}) = exp(-\| \mathbf{k} \|_\alpha^\alpha), \tag{14}$$

where $\|\mathbf{a}\|_\alpha = \left(\sum_{j=1}^n |a_j|^\alpha \right)$ denotes the l_α norm.

If the characteristic function of $\mathbf{X}$ is of the form 14 we say that $\mathbf{X}$ possesses an *α-symmetric multivariate distribution* [3]. For $\alpha = 2$, the 2-symmetric multivariate distribution reduces to a spherical distribution. In other cases ($\alpha < 2$) the symmetric multivariate distribution is only invariant under the group of permutations. Let $\mathcal{P}$ be the permutation group, i.e. if $\mathbf{H} \in \mathcal{P}$, then $\mathbf{H}^T\mathbf{H} = \mathbf{I}$ and the elements of $\mathbf{H}$ are only 0 or 1. If $\mathbf{X} \in \mathbf{S_\alpha S}$ then $\mathbf{HX} \sim \mathbf{S_\alpha S}$. Figure 7 presents selected 2-D pdfs of α-symmetric multivariate distribution. The lack of spherical symmetry influences the relation between the effectiveness of an EA in a multimodal optimization task an a reference frame selection. This fact, called the *symmetry effect*, was studied by Obuchowicz [14], [15], who analyzed the non-spherical Cauchy mutation applied in the ESSS and Evolutionary Programming (EP) algorithms.

Due to the fact that multivariate α-stable mutation is a biased genetic operator, it is obvious that effectiveness of $ESSS_\alpha$ algorithm for multidimensional environment must differs from results presented in this paper. Nevertheless, the results obtained for one-dimensional α-stable mutations cast light on the issue of using their multivariate spherically symmetrical counterparts. Above affirmation seems to be justified by Rudolphs analytic analyzes [20], who proved that order of local convergence is identical for Gaussian and spherical Cauchy distribution.

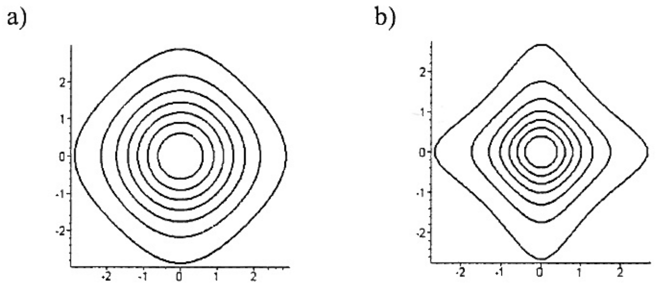

Fig. 7. Contour maps of 2-D probability density functions of $S_{\alpha}S$ for $\alpha = 1.5$ (a) and $\alpha = 1$ (b)

VIII. Conclusion

The most challenging problem of global optimization is how to establish the compromise between two mutually exclusive property: accuracy in locating potential solution and ability of escaping from local optima. The results presented in this work revel very interesting relationship between parameters of stable distributions, and their influence on the global optimization effectiveness. The family of Symmetric Stable Distribution $S\alpha S(\sigma)$ is characterized by two parameters: stable index α and scale σ. To investigate effectiveness of $ESSS_{\alpha}$ algorithm in global optimization problems, series of simple experiments have been done. The domain of the parameters has been sampled in a way to gain an insight into problem of properly choice random distribution to the mutate operator. Among many tested parameters one can find two of them, which correspond to the very often use distribution: Gaussian and Cauchy. It makes an opportunity to compare efficacy of well know mutation operators with other belonging to the same stable mutation's family.

Taking into consideration nature of mutation, one can distinguish between micro-mutations, which improve accuracy of finding solution, and macro-mutations which make possibility of escaping from local optima. It is worthy to be stressed, that mutate operator generate some individuals which are neutral to the process of global optimization i.e. they do not improve neither local nor global convergence. Thereby number of useless steps can be minimized by properly chose of stable mutations parameters.

From authors earlier research [17] follows that results concerning α-stable distributions obtained for one-dimensional search space can not be simple generalized on the high dimensional cases. The main obstacle is connected with the fact that multidimensional stable distribution possess heterogenous symmetry. To avoid this awkward problem authors want to consider family of stable spherically symmetry distribution in mutation operator. It will be motivation to the authors further work.

References

[1] T. Bäck and H.-P. Schwefel, "An overview of evolutionary computation," *Evol. Comput.* Vol. 1, No 1, pp.1–23, 1993

[2] T. Bäck, D.B. Fogel and Z. Michalewicz, (Eds.) *Handbook of Evolutionary Computation*, Institute of Physics Publishing and Oxford University Press, NY, 1997

[3] K.-T. Fang, S. Kotz, and K.-W. Ng, *Symmetric Multivariate and Related Distributions*. Chapman and Hall, London, 1990.

[4] L.J. Fogel, A.J. Owens and M.J. Walsh, *Artificial Intelligence through Simulated Evolution*, Wiley, New York, 1966.

[5] D.B. Fogel, "An introduction to simulated evolutionary computation," *IEEE Trans. Neural Networks*, Vol. 5, pp.3–14, 1994

[6] R. Galar, "Handicapped individua in evolutionary processes", *Biological Cybernetics*, Vol.51, 1985, pp.1–9.

[7] R. Galar, "Evolutionary search with soft selection", *Biological Cybernetics*, Vol.60, 1989, pp.357–364.

[8] M. Gutowski, "Lévy flights as an underlying mechanism for a global optimization algorithm", *Proc. 5th Conf. Evolutionary Algorithms and Global Optimization*, Jastrzębia Góra, Poland, Warsaw University of Technology Press, 2001, pp. 79–86.

[9] I. Karcz-Dulęba, "Dynamics of infinite populations envolving in a landscape of uni- and bimodal fitness functions", *IEEE Trans. Evolutionary Computation*, Vol.5, No.4, 2001, pp.398–409.

[10] P. Lévy, *Calcul des Probabilites*, Gauthier Villars, Paris, 1925.

[11] Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs*, Springer–Verlag, Berlin Heidelberg, 1996.

[12] J.P. Nolan, *Stable Distributions, Models for Heavy Tailed Data*, Springer–Verlag, Berlin Heidelberg, 2003

[13] A. Obuchowicz, "On the true nature of the multi-dimensional Gaussian mutation", In: V. Kurkova, N.C. Steel, R. Neruda and M. Karny (Eds.) *Artificial Neural Networks and Genetic Algorithms*, Springer, Vienna 2001, pp.248–251.

[14] A. Obuchowicz, "Mutli-dimensional Gaussian and Cauchy mutations", in: M. Kłopotek, M. Michalewicz, and S.T. Wierzchoń (Eds.) *Intelligent Information Systems*, Physica–Verlag, Heidelberg, 2001, pp. 133–142.

[15] A. Obuchowicz, "Multidimensional mutations in evolutionary algorithms based on real-valued representation", *International Journal of System Science*, Vol.34, No.7, 2003, pp.469–483.

[16] A. Obuchowicz, *Evolutionary Algorithms in Global Optimization and Dynamic System Diagnosis*, Lubuskie Scientific Society, Zielona Góra 2003.

[17] A. Obuchiwicz and P.Prętki, *Analysis of phenotypic evolution with symmetric α-stable mutation*, Proc. 7th Conf. on Evolutionary Computation and Global Optimization, Kazimierz Dolny,Poland, Warsaw University of Technology Press, 2004, pp. 123–135.

[18] I. Rechenberg, *Cybernetic solution path of an experimental problem*, Roy. Aircr. Establ., libr. Transl. 1122, Farnborough, Hants., UK, 1965.

[19] I. Rechenberg, *Ewolutionsstrategie: Optimierung technicher Systeme nach Principien der biologischen Evolution.*, Frommann–Holtboog, Stuttgart, 1973.

[20] G. Rudolph, "Local convergence rates of simple evolutionary algorithms with Cauchy mutations", *IEEE Trans. Evolutionary Computation*, Vol.1, No.4, pp.249–258, 1997.

[21] H.-P., Schwefel, *Numerical Optimization of Computer Models.*, Wiley, Chichester, 1981.

[22] G. Samorodnitsky and M.S. Taqqu, *Stable Non-Gaussian Random Processes*, Chapman & Hall, New York, 1994

[23] R.Weron, *Correction to: On the Chambers-Mallows-Stuck method for simulating skewed stable random variables*, Research Report, Wrocław University of Technology, 1996

[24] R. Weron, "Lévy-stable distributions revisited: tail index > 2 does not exclude the Lévy-stable regime, *Int. Journ. Modern Physics C*, Vol.12, No.2, pp.209–223, 2001.

[25] A. Zolotariev, *One-Dimensional Stable Distributions*, American Mathematical Society, Providence, 1986.