**Mathematical modeling**

**Математическое моделирование**

*UDC 004.023, 519.677*
*https://doi.org/10.32362/2500-316X-2026-14-1-103-112*
*EDN LDJQIL*



RESEARCH ARTICLE

# Experimental investigation of convergence characteristics of quasi-Newton algorithm on nonsmooth and nonconvex functions

**Alexander V. Smirnov** @

*MIREA – Russian Technological University, Moscow, 119454 Russia*
@ *Corresponding author, e-mail: av_smirnov@mirea.ru*

• **Submitted:** 02.04.2025 • **Revised:** 14.06.2025 • **Accepted:** 13.11.2025

**Abstract**

**Objectives.** The aim of the paper is to develop a methodology for studying the convergence of the quasi-Newton minimization algorithm (QNA) on nonsmooth and nonconvex objective functions (OF), as well as to conduct related numerical experiments.

**Methods.** The experiments were performed on a flexible OF capable of mimicking various patterns of value changes in different directions away from the minimum. A total of 18 OF instances with different landscape parameters were studied. For each example, 200 QNA searches were performed from random staring points, and all corresponding OF values were recorded. Then, the Expected Run Time (ERT) to reach a given threshold level of the OF was computed based on the data. The dependence of the achieved OF threshold on ERT was approximated separately for the segment in which all thresholds were achieved in all searches, and for a segment in which the thresholds were achieved, but not in all searches.

**Results.** The experiments show that, for the majority of cases in which all thresholds are achieved in all takes, a decrease in the OF follows the geometric progression law (linear convergence). However, in the second segment, convergence follows the power law. It was also found that the presence of anisotropy of the OF landscape and a loss of smoothness lead to convergence slowdown, and premature termination of search process before reaching the minimum with the required accuracy.

**Conclusions.** The study identifies patterns in the QNA convergence on the objective functions with different landscape parameters. Further advancement of the methodology would involve automating data collection and processing, as well as extending it to other types of optimization algorithms.

**Keywords:** quasi-Newton algorithm, objective function landscape, convex function, concave function, nonsmooth function, approximation, exponent, algorithm convergence

**For citation:** Smirnov A.V. Experimental investigation of convergence characteristics of quasi-Newton algorithm on nonsmooth and nonconvex functions. *Russian Technological Journal.* 2026;14(1):103–112. https://doi.org/10.32362/2500-316X-2026-14-1-103-112, https://www.elibrary.ru/LDJQIL

**Financial disclosure:** The author has no financial or proprietary interest in any material or method mentioned.

The author declares no conflicts of interest.

НАУЧНАЯ СТАТЬЯ

# Экспериментальное исследование характеристик сходимости квазиньютоновского алгоритма на негладких и невыпуклых функциях

**А.В. Смирнов** [@]

*МИРЭА – Российский технологический университет, Москва, 119454 Россия*
[@] *Автор для переписки, e-mail: av_smirnov@mirea.ru*

• **Поступила:** 02.04.2025 • **Доработана:** 14.06.2025 • **Принята к опубликованию:** 13.11.2025

**Резюме**

**Цели.** Целью работы является разработка методики исследования характеристик сходимости квазиньютоновского алгоритма (КНА) на негладких и невыпуклых целевых функциях (ЦФ) и выполнение экспериментов по этой методике.

**Методы.** Эксперименты выполнялись на тестовой функции, обеспечивающей возможность задания различных законов изменения ее значений по разным направлениям от точки минимума. Всего исследованы 18 примеров ЦФ с разными параметрами рельефа. Для каждого примера выполнялись 200 стартов КНА из случайных точек и фиксировались все значения ЦФ, полученные в процессе поиска. Затем по этим данным вычислялись значения Expected Run Time (ERT) – ожидаемого времени достижения заданного порогового уровня ЦФ. Далее выполнялась аппроксимация зависимости достигнутого порога ЦФ от ERT отдельно для отрезка, в котором все пороги достигаются во всех стартах для этого примера, и для отрезка, в котором пороги достигаются, но не во всех стартах.

**Результаты.** Эксперименты показали, что в большинстве примеров для отрезка, в котором все пороги достигаются во всех стартах, имеет место убывание ЦФ по закону геометрической прогрессии (линейная сходимость), а во втором отрезке преобладает сходимость по степенному закону. Также установлено, что наличие анизотропии рельефа ЦФ и нарушений гладкости приводят к замедлению сходимости и завершению поиска до достижения минимума с требуемой точностью.

**Выводы.** Исследование позволило выявить закономерности в сходимости КНА на ЦФ с различными свойствами рельефа. Дальнейшее развитие методики должно включать автоматизацию сбора и обработки данных и распространение на другие виды алгоритмов поиска оптимальных решений.

**Ключевые слова:** квазиньютоновский алгоритм, рельеф целевой функции, выпуклая функция, вогнутая функция, негладкая функция, аппроксимация, показатель степени, сходимость алгоритма

**Для цитирования:** Смирнов А.В. Экспериментальное исследование характеристик сходимости квазиньютоновского алгоритма на негладких и невыпуклых функциях. *Russian Technological Journal*. 2026;14(1):103–112. https://doi.org/10.32362/2500-316X-2026-14-1-103-112, https://www.elibrary.ru/LDJQIL

**Прозрачность финансовой деятельности:** Автор не имеет финансовой заинтересованности в представленных материалах или методах.

Автор заявляет об отсутствии конфликта интересов.

## INTRODUCTION

The paper considers the problem of finding the local extremum (LE) (minimum) $\mathbf{x}^*$ of the objective function (OF) $f(\mathbf{x})$ within a given search space $\Omega_X$:

$$\mathbf{x}^* = \arg\min_{\mathbf{x}\in\Omega_X}(f(\mathbf{x})). \quad (1)$$

Quasi-Newton algorithms (QNA) can be an effective method for solving this problem. These algorithms use an approximation of the Hessian matrix (the matrix of second partial derivatives) based on the OF gradient which is only the first derivative. This allows the algorithm to determine the direction of the next step in the search process. It has been shown that for these algorithms to converge to the minimum, the function must be smooth and convex within the search space $\Omega_X$ [1, 2]. Software QNA implementations are available in widely used mathematical programming packages.

In cases where the conditions for smoothness and/or convexity of the objective function are not met, a rigorous analysis of the QNA is not available. Nevertheless, in practice, QNA has been successfully used to find local extrema of nonsmooth and/or nonconvex OFs in many cases. The conditions under which this application is successful, and the achievable results, have been little explored. At the same time, studies on the convergence conditions and rates of certain other search algorithms at such OFs are available [1–7]. However, these algorithms tend to converge more slowly than QNA.

The study [8] presents a theoretical analysis of the QNA performance for the function of a single variable, $f(x) = |x|$, with a loss of smoothness at the LE point. The analysis demonstrates that, in order to achieve the LE with an accuracy of no greater than a specified $\varepsilon > 0$, a logarithmic number of iterations ($\log_2(\varepsilon^{-1})$) is required. It has not been possible to perform theoretical analysis for the function $f(\mathbf{x}) = ||\mathbf{x}||$ with a search space dimension $ND > 1$. However, experimental results show that there is a linear convergence in the search process, with the OF value decreasing following the geometric progression law, in which the denominator approaches 1 as $ND$ increases. The study uses only a specific type of the objective function with fixed landscape parameters, and results for other OF types may vary.

The study [9] presents the results of experimental investigations on QNA and a hybrid algorithm based on it, using a large range of test functions. However, the paper lacks sufficient information regarding the convergence characteristics, since only the final outcomes are provided: namely, the achieved LE values and the number of iterations. Additionally, there is a lack of information regarding the relationship between the number of iterations and the variation in the OF value within the LE.

A comparison of the performance of several algorithms, including QNA implementation, is conducted in [10]. This study utilizes a basic set of testing functions [11] and statistical indicators of convergence processes. However, this selection of testing functions does not permit a wide variation in the characteristics of the OF landscape in the LE vicinity, particularly with regard to convexity and asymmetry. In [12], experiments are conducted using QNA, and a testing function with the capability of adjusting certain features. However, convergence processes are not examined, and only the achieved OF values are recorded. Furthermore, it is possible to set only two distinct laws for modifying the OF during experiments in coordinates.

This study aims to develop a methodology and conduct experimental investigations into the convergence of QNA at various OF landscape parameters in the LE vicinity, including loss of smoothness, nonconvexity, anisotropy, and asymmetry across coordinates.

## RESEARCH METHODS AND MATERIALS

Experimental research was conducted using the *MATLAB* software[1] implemented on the *GNU Octave* freeware platform[2]. The quasi-Newton algorithm within this software was implemented using the fminunc(..) function. However, this function only returns the final search result and the total number of the OF evaluations. In order to analyze convergence characteristics, all OF values calculated during the search process need to be obtained. In order to address this, a new function called QNLS_M(..) was developed. The simplified structure is shown in Fig. 1.

1. Input: **X** is a starting point; *Options* is the algorithm configuration.
2. Initialization of the Hessian **B** approximation by a diagonal matrix.
3. Calculation of the objective function *F* and its gradient **g** at the starting point.
4. Iterate the search until the completion condition is met.
   4.1. Determine the search direction, **s**, based on **B** and **g**.
   4.2. Find the appropriate step size, λ, using the linear search method.
   4.3. Move to the new point **X** = **X** + λ**s** and calculate the *F* and **g** values at this point.
   4.4. Calculate the **B** value using the BFGS[3] algorithm.
   4.5. Verify search completion conditions.
5. Output: **History** is the array containing the calculated OF values during the search in the order they were calculated, as well as the coordinates of the points where the calculations were performed.

**Fig. 1.** Structure of the QNLS_M (..) function implementing QNA

[1] https://www.mathworks.com/products/matlab.html. Accessed August 26, 2025.

[2] https://octave.org/. Accessed August 26, 2025.

[3] The Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm is an iterative method for solving unconstrained nonlinear optimization problems.

This structure is standard, except for item 5 [1, 2]. In item 4, the linear search algorithm described in [8] is used. In item 4.4, a new approximation of the Hessian **B** is calculated when the condition for positive curvature of the objective function is met at this iteration. The finite difference method is used to calculate the gradient **g** and derivative in the direction of the linear search. In order to minimize the number of calculations of the OF, only one additional point with a bias $\Delta x_i = 10^{-8}$ is taken to estimate the partial derivative for each $x_i$ coordinate.

Table 1 shows the search completion conditions, indicator values for the reasons for ExitFlag completion, and thresholds used in the experiments.

**Table 1.** Indicator values for the reasons of search completion

| ExitFlag | Search completion condition |
|---|---|
| 0 | Maximum iteration limit of max_iter = 1000 is exceeded |
| 1 | First-order optimality criterion $\max(\|\mathbf{g}\|) \le$ g_tol $= 10^{-8}$ is met |
| 2 | Step size condition $\|\lambda\mathbf{s}\| \le$ min_step $= 10^{-12}$ is met |
| 3 | The number of consecutive iterations with a relative change in the OF less than ftol = 1.00001 has exceeded the specified threshold max_Nftol = 4 |
| 4 | No step is found that does not result in an increase in the OF |
| 5 | The OF value is less than the specified threshold FuncTol $= 10^{-8}$ |

In the experiments, the test function TestLE6(**x**, **x***, **R**, **W**, **K**, *NS*) is used. The value at point **x** is calculated as in the following way. First, the bias from the given LE **x*** and the rotation of the coordinates given by the matrix **R** are performed:

$$\mathbf{z} = (\mathbf{x} - \mathbf{x}^*)^{\mathrm{T}}\,\mathbf{R}, \tag{2}$$

wherein T represents the transpose operation. Then, for coordinates with indices ranging from 1 to $ND - NS$ (where $NS$ is an integer), the interpolation of the $k$ coefficients and the exponent α is carried out, along with the calculation of a preliminary value of the function $f$:

$$k = \frac{1}{\|\mathbf{z}\|^2}\sum_{n=1}^{ND-NS}(K_{1n}z_n{}^2h(z_n) + K_{2n}z_n{}^2h(-z_n)),$$

$$\text{wherein } h(y) = \begin{cases} 1, y > 0, \\ 0, y \le 0, \end{cases} \tag{3}$$

$$\alpha = \frac{1}{\|\mathbf{z}\|^2}\sum_{n=1}^{ND-NS}(W_{1n}z_n{}^2h(z_n) + W_{2n}z_n{}^2h(-z_n)),$$

$$f = k\|\mathbf{z}\|^{\alpha}.$$

The variables $K_{ij}$ and $W_{ij}$ are elements of the matrices **K** and **W** which have dimensions of $2 \times ND$. They represent the values of coefficients and exponents, respectively, in the positive and negative directions for all coordinates in the search space. This algorithm, proposed in [13], allows for arbitrary setting of power function parameters for different coordinates and for smooth changes in these parameters in intermediate directions.

If $NS = 0$, then the final value of the function $f$ is obtained. If $NS > 0$, a computation cycle is carried out for coordinates with indices $n$ ranging from $ND - |NS| + 1$ to $ND$ as follows:

$$f = f + \\ + \max\left[K_{1n}|z(n)|^{W_{1n}}\operatorname{sign}(z(n)), -K_{2n}|z(n)|^{W_{2n}}\operatorname{sign}(z(n))\right]. \tag{4}$$

If $NS < 0$, a series of operations are carried out at the specified indices, as follows:

$$f = \\ = \max\left[f, K_{1n}|z(n)|^{W_{1n}}\operatorname{sign}(z(n)), -K_{2n}|z(n)|^{W_{2n}}\operatorname{sign}(z(n))\right]. \tag{5}$$

As a result, lines originating from the LE point are generated, along which smoothness of the function is lost. For $NS = 0$, the smoothness can only be lost at the LE point **x***. By varying the combination of parameters, many well-known unimodal OFs, as well as functions with novel properties can be generated. Examples of graphs of the TestLE6(..) function for $ND = 2$ and various $NS$ values are presented in Fig. 2. The other parameters of the function have the following values:

$$\mathbf{x}^* = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \mathbf{W} = \begin{pmatrix} 1.5 & 0.5 \\ 0.75 & 1.0 \end{pmatrix}, \quad \mathbf{K} = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}.$$

Experimental studies were conducted in the form of individual tests, each consisting of 10 sets of 20 local extremum searches. The initial search points were randomly selected on a hypersphere with unit radius and center at **x***. In each test, various combinations of the **W**, **K**, and $NS$ parameters were applied. At the start of each set of searches, the coordinate axes were rotated using a randomly generated and orthonormalized matrix **R**. During each search, the coordinates of all points were recorded where the OF calculation was performed, as well as the actual OF values.

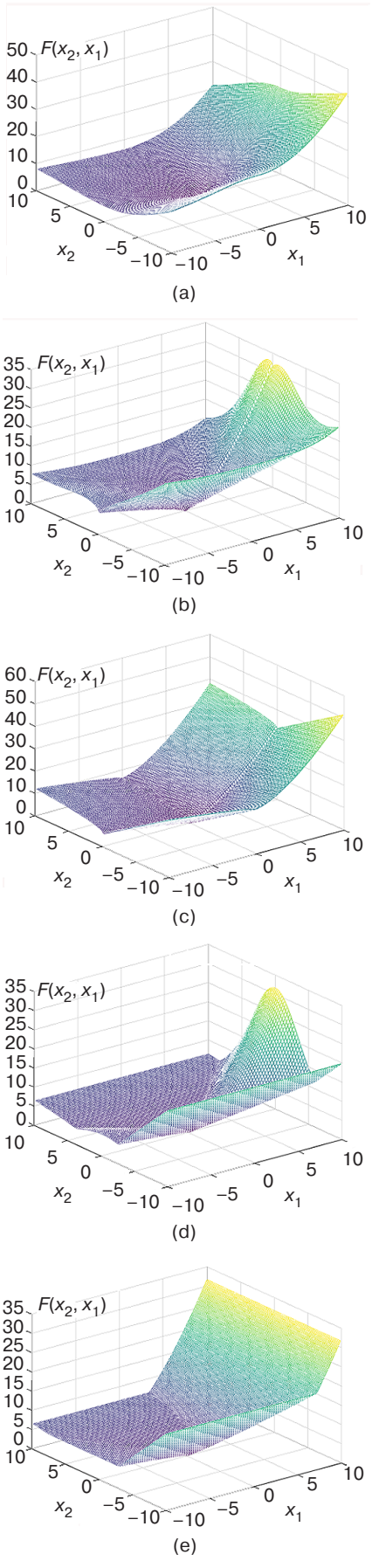


**Fig. 2.** Graphs of the TestLE6(..) function:
(a) *NS* = 0; (b) *NS* = 1; (c) *NS* = 2; (d) *NS* = −1; and (e) *NS* = −2

The expected time to reach a certain OF threshold was used as a characteristic to summarize the results of all tests. This is known in optimization literature as the Expected Run Time (ERT). This function can be defined as follows [14, 15].

Let us introduce a variable:

$$h = -\lg(f), \tag{6}$$

wherein $f$ is the OF value. We define the scale of the threshold levels, as follows:

$$h_k = -2 + 0.02(k-1), \quad k = \overline{1,51}. \tag{7}$$

The corresponding OF thresholds, $f_k$, derived from (6), range from $10^2$ to $10^{-8}$. We define in the following way: $Nr_i(h_k)$ as the number of the OF computations performed during the $i$th search until threshold $h_k$ is reached; and $Ne_i$ as the total number of the OF computations until the $i$th search completes. $Ir_k$ represents a subset of searches which reach the threshold $h_k$, with $Inr_k$ denoting the number of these searches. $Inr_k$ denotes a subset of searches that do not reach the threshold $h_k$. Then,

$$\mathrm{ERT}(h_k) = \frac{\sum\limits_{i \in Ir_k} Nr_i(h_k) + \sum\limits_{i \in Inr_k} Ne_i}{|\,Ir_k\,|}. \tag{8}$$

If the threshold $h_k$ is not reached in any of the searches, the ERT($h_k$) value becomes infinitely large.

We introduce the indicators determined by ERT, in order to compare the characteristics of the QNA operation. For this purpose, we define the following boundaries for the threshold scale:

$h_{\mathrm{start}}$ is the threshold value reached at the beginning of the experiment;

$h_{\mathrm{all}}$ is the maximum threshold value reached during the experiment;

$h_{\mathrm{finish}}$ is the highest threshold value achieved in at least one search during the experiment.

If $h_{\mathrm{all}} = 8$, which is the upper limit of the scale, then the value $h_{\mathrm{finish}} = h_{\mathrm{all}}$ is set.

In the segment $[h_{\mathrm{start}}, h_{\mathrm{all}}]$, we approximate the relationship between lg($f$) and ERT using the *MATLAB/Octave regress*(..) function, as follows:

$$\lg(f) = b_{11} \cdot \mathrm{ERT} + b_{12} \cdot \lg(\mathrm{ERT}) + b_{13}. \tag{9}$$

We also calculate the determination coefficients for the first two terms.

If the segment $[h_{\mathrm{all}}, h_{\mathrm{finish}}]$ contains between 2 and 9 divisions on the threshold scale, then we perform approximation of the form (9) for the entire segment. We also define the coefficients $b_{31}$, $b_{32}$, and $b_{33}$, as well as the corresponding determination coefficients.

If this segment spans at least 9 divisions, it is divided into two subsegments, with the boundary of the second subsegment on the threshold scale denoted by $h_b$. Approximation of the form (9) is performed separately for each of these subsegments. In this case, coefficients for the first subsegment are denoted as $b_{21}$, $b_{22}$, and $b_{23}$, while coefficients for the second subsegment are $b_{31}$, $b_{32}$, and $b_{33}$. The value of $h_b$ is determined by iterating through possible values within the $[h_{all}, h_{finish}]$ segment. Approximations for each subsegment are calculated and the value that minimizes the mean-square error of approximation is selected. Splitting the $[h_{all}, h_{finish}]$ segment allows for experiments to obtain more accurate approximations of the dependence of lg($f$) on ERT in that segment.

Subsequently, by applying potentiation (9), we derive the following

$$f = \left(10^{b_{11}}\right)^{\mathrm{ERT}} \mathrm{ERT}^{b_{12}} \cdot 10^{b_{13}}. \qquad (10)$$

We denote $10^{b_{11}} = q_1$, $b_{12} = p_1$, $10^{b_{13}} = a_1$. Then (10) takes the following form:

$$f = a_1 \cdot q_1^{\mathrm{ERT}} \cdot \mathrm{ERT}^{p_1}. \qquad (11)$$

The second cofactor in (11) varies in accordance with the law of geometric progression, corresponding to linear convergence [1, 4]. The third cofactor follows a power function. Therefore, the values of $q_1$ and $p_1$ determine the laws and rates of convergence for the search on the segment $[h_{start}, h_{all}]$. Likewise, the values of $q_3$ and $p_3$ for the segment $[h_{all}, h_{finish}]$ can be determined if this interval is not divided; or the values of $q_2$ and $p_2$ for $[h_{all}, h_b]$, and the values of $q_3$ and $p_3$ for $[h_b, h_{finish}]$, if the interval $[h_{all}, h_{finish}]$ is split into two.

Segment boundaries and approximation coefficients for these segments can serve as indicators to compare the convergence of search algorithms for OF with different properties. In this method, instead of the OF value, the distance from the current point **x** to LE $\mathbf{x}^*$ can be used as an indicator of convergence [1, 5, 6]. The ERT($f$) dependence can also be approximated, in order to obtain estimates for the number of the OF calculations required until a specified OF value is reached. This is another useful characteristic in certain cases [4, 7].

## EXPERIMENTAL RESULTS

The parameters of the TestLE6(**x**, $\mathbf{x}^*$, **R**, **W**, **K**, *NS*) function, as determined in 18 tests, are presented in Table 2. The LE location is always set to the origin, $\mathbf{x}^* = \mathbf{0}$.

In tests 1–6, the space dimension $ND = 4$ and the OF is isotropic in all directions. In tests 1–3, the OF is convex. In test 4, it changes according to a linear law, and in tests 5 and 6, it is concave. In test 1, the objective function is smooth everywhere, while in tests 2–5, there is a loss of smoothness in the LE. Subsequent tests use the OFs from tests 2 and 5 as a basis, and modify any properties. Tests 7–10 introduce anisotropy in certain directions. Tests 11–14 increase the dimension of the space, while tests 15–18 create loss of smoothness on lines originating from the LE.

Let us now explain the introduction of anisotropy. The calculation procedure for **W** elements is as follows:

**Table 2.** Test function parameters in experiments

| Test | *ND* | $W_{max}$ | $W_{min}$ | $K_{max}$ | $K_{min}$ | *NS* | Test | *ND* | $W_{max}$ | $W_{min}$ | $K_{max}$ | $K_{min}$ | *NS* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 4 | 2.00 | 2.00 | 1 | 1 | 0 | 10 | 4 | 0.75 | 0.75 | 1 | 0.1 | 0 |
| 2 | 4 | 1.50 | 1.50 | 1 | 1 | 0 | 11 | 8 | 1.50 | 1.50 | 1 | 1 | 0 |
| 3 | 4 | 1.25 | 1.25 | 1 | 1 | 0 | 12 | 16 | 1.50 | 1.50 | 1 | 1 | 0 |
| 4 | 4 | 1.00 | 1.00 | 1 | 1 | 0 | 13 | 8 | 0.75 | 0.75 | 1 | 1 | 0 |
| 5 | 4 | 0.75 | 0.75 | 1 | 1 | 0 | 14 | 16 | 0.75 | 0.75 | 1 | 1 | 0 |
| 6 | 4 | 0.50 | 0.50 | 1 | 1 | 0 | 15 | 4 | 1.50 | 1.50 | 1 | 1 | 2 |
| 7 | 4 | 1.50 | 1.50 | 10 | 1 | 0 | 16 | 4 | 1.50 | 1.50 | 1 | 1 | –2 |
| 8 | 4 | 1.50 | 0.75 | 1 | 1 | 0 | 17 | 4 | 0.75 | 0.75 | 1 | 1 | 2 |
| 9 | 4 | 1.50* | 0.75* | 1 | 1 | 0 | 18 | 4 | 0.75 | 0.75 | 1 | 1 | –2 |

$$W_{1i} = W_{2i} = W_{\min} + \frac{(W_{\max} - W_{\min})(i-1)}{ND-1},\ i = \overline{1,ND}. \quad (12)$$

For $W_{\max} = W_{\min}$, we obtain $W_{1i} = W_{2i} = W_{\max}$, $i = \overline{1,ND}$, and the exponent of α in (5) is independent of direction. Similarly, values for the matrix **K** elements were determined. In test 9, a random permutation of the **W** row elements was also performed, for example:

$$W(\text{Test } 8) = \begin{pmatrix} 0.75 & 1 & 1.25 & 1.5 \\ 0.75 & 1 & 1.25 & 1.5 \end{pmatrix},$$

$$W(\text{Test } 9) = \begin{pmatrix} 1 & 1.25 & 1.5 & 0.75 \\ 0.75 & 1.5 & 1 & 1.25 \end{pmatrix}.$$

Table 3 shows the reasons for completion of the search in the experiments (Table 1). In the table, $EF_1$ (Exit Flag) represents the most frequently encountered reason in the certain test, while $PEF_1$ indicates the probability of that reason. Additionally, other reasons for termination encountered in the experiment, ranked in descending order by probability, are provided. This data indicates that if the function is convex at least along part of coordinates, i.e., if there is at least one element of the matrix **W** exceeding one, the search will reach a value of $f = 10^{-8}$. Otherwise, the search process terminates prematurely for other reasons.

Figures 3 and 4 show the ERT graphs obtained from the experiments, while Table 4 presents the approximation results for these dependencies. In order to enhance clarity, thresholds $h_k$ are not used as the argument, but rather the corresponding $f$ values obtained using (6), with the same index, for example, $f_{\text{all}} = 10^{-h\text{all}}$. The $\lg(f_{\text{start}})$ values are not included in the table. They can be inferred from the point at which the ERT increases. The approximation coefficients are presented in $P/R^2$ format, where $P$ represents a value of $q$ or $p$, and $R^2$ represents the corresponding determination coefficient. Only points where the ERT values are finite are shown in the graphs.

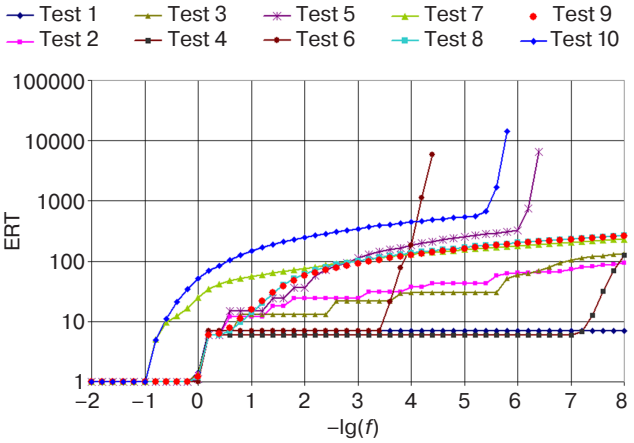


**Fig. 3.** ERT experiment graphs: tests 1–10

Let us proceed to analyzing the results of the experiments. For the isotropic quadratic function in test 1, the QNA algorithm accurately identifies the LE after the first iteration, illustrated by the horizontal line in Fig. 3. However, the approximation of the form (9) is highly inaccurate. The corresponding values in Table 4 are highlighted in bold.

**Table 3.** Statistics of search completion reasons in experiments

| Test | $EF_1$ | $PEF_1$ | $EF_2$ | $PEF_2$ | Test | $EF_1$ | $PEF_1$ | $EF_2$ | $PEF_2$ | $EF_3$ | $PEF_3$ | $EF_4$ | $PEF_4$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 5 | 0.970 | 1 | 0.030 | 10 | 2 | 1.000 | – | – | – | – | – | – |
| 2 | 5 | 1.000 | – | – | 11 | 5 | 1.000 | – | – | – | – | – | – |
| 3 | 5 | 1.000 | – | – | 12 | 5 | 1.000 | – | – | – | – | – | – |
| 4 | 2 | 0.530 | 5 | 0.470 | 13 | 2 | 1.000 | – | – | – | – | – | – |
| 5 | 2 | 1.000 | – | – | 14 | 2 | 1.000 | – | – | – | – | – | – |
| 6 | 2 | 0.990 | 4 | 0.010 | 15 | 5 | 0.650 | 2 | 0.340 | 3 | 0.010 | – | – |
| 7 | 5 | 1.000 | – | – | 16 | 5 | 0.685 | 2 | 0.155 | 3 | 0.140 | 4 | 0.020 |
| 8 | 5 | 1.000 | – | – | 17 | 2 | 0.895 | 4 | 0.080 | 3 | 0.025 | – | – |
| 9 | 5 | 1.000 | – | – | 18 | 2 | 0.450 | 4 | 0.300 | 3 | 0.250 | – | – |

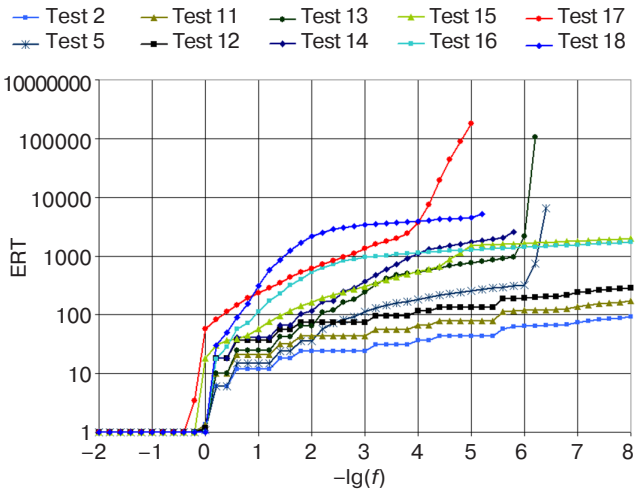


**Fig. 4.** ERT experiment graphs: tests 2, 5, and 11–18

In tests 4 and 6, with OF exponents of 1 and 0.5, respectively, the QNA also jumped to points with $f = f_{all}$ after the first iteration. On the $[f_{all}, f_{finish}]$ segments, the search slowed down and these sections can be approximated by a power law with good accuracy. There is no descent by geometric progression because $q_3 = 1.00$. In tests 2, 3, and 5 with exponents of 1.5, 1.25, and 0.75, respectively, no such jump in OF was observed. The $[f_{start}, f_{all}]$ segment in tests 2 and 5 can be accurately approximated by geometric progression.

In subsequent experiments, the OF parameters from tests 2 and 5 were altered. As can be seen from the graphs in Fig. 3, introducing the form (5) anisotropy into the OF, either by coefficient (tests 7 and 10) or by exponent (tests 8 and 9), resulted in an upward shift in ERT, i.e., a slowing down of the search. At the same time, random permutations of the degree indicators across the coordinates in test 9 did not alter ERT compared to test 8. In all experiments, the $[f_{start}, f_{all}]$ segment is approximated with high precision using

**Table 4.** ERT approximation results in experiments

| Test | $\lg f_{all}$ | $\lg f_b$ | $\lg f_{finish}$ | ERT($f_{all}$) | ERT($f_b$) | ERT($f_{all}$) | $q_1$ | $p_1$ | $q_2$ | $p_2$ | $q_3$ | $p_3$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | −8.0 | – | −8.0 | 7.00 | – | – | **0.15/0.14** | **0.74/0.14** | – | – | – | – |
| 2 | −8.0 | – | −8.0 | 93.2 | – | – | 0.84/0.96 | −1.06/0.78 | – | – | – | – |
| 3 | −8.0 | – | −8.0 | 134 | – | – | 0.96/0.78 | −3.39/0.87 | – | – | – | – |
| 4 | −7.2 | – | −8.0 | 6.77 | – | 126 | **0.00/0.19** | **14.0/0.17** | – | – | 1.00/0.94 | −0.56/1.00 |
| 5 | −5.8 | – | −6.4 | 308 | – | $6.56 \cdot 10^3$ | 0.97/0.96 | −0.97/0.87 | – | – | 1.00/0.79 | −0.58/0.94 |
| 6 | −3.4 | – | −4.4 | 7.00 | – | $5.86 \cdot 10^3$ | **0.00/0.30** | **66.7/0.29** | – | – | 1.00/0.70 | −0.33/0.98 |
| 7 | −8.0 | – | −8.0 | 232 | – | – | 0.91/1.00 | 0.057/0.72 | – | – | – | – |
| 8 | −8.0 | – | −8.0 | 270 | – | – | 0.94/1.00 | −0.11/0.75 | – | – | – | – |
| 9 | −8.0 | – | −8.0 | 263 | – | – | 0.94/1.00 | −0.28/0.76 | – | – | – | – |
| 10 | −5.2 | – | −5.8 | 563 | – | $14.3 \cdot 10^3$ | 0.98/1.00 | −0.26/0.73 | – | – | 1.00/0.80 | −0.58/0.95 |
| 11 | −8.0 | – | −8.0 | 170 | – | – | 0.91/0.96 | −0.84/0.74 | – | – | – | – |
| 12 | −8.0 | – | −8.0 | 290 | – | – | 0.94/0.96 | −0.60/0.68 | – | – | – | – |
| 13 | −5.8 | – | −6.2 | 951 | – | $106 \cdot 10^3$ | 0.99/0.92 | −1.01/0.88 | – | – | – | – |
| 14 | −4 | −5.4 | −5.8 | 1.11k | $1.96 \cdot 10^3$ | $2.53 \cdot 10^3$ | 1.00/0.75 | −1.04/0.88 | 1.00/0.99 | −6.66/1.00 | 1.02/0.89 | −52.4/0.90 |
| 15 | −4.2 | −5.2 | −8.0 | 602 | $1.53 \cdot 10^3$ | $1.98 \cdot 10^3$ | 0.94/0.99 | −0.39/0.83 | 1.07/0.86 | −14.8/0.90 | 0.75/0.99 | 64.0/0.99 |
| 16 | −0.8 | −3.6 | −8.0 | 71.9 | $1.09 \cdot 10^3$ | $1.73 \cdot 10^3$ | 0.98/0.98 | −0.064/0.80 | 0.99/0.98 | 0.19/0.91 | 0.99/1.00 | −11.4/1.00 |
| 17 | −0.6 | −3.8 | −5.0 | 140 | $2.45 \cdot 10^3$ | $177 \cdot 10^3$ | 0.99/0.97 | −0.33/0.84 | 1.00/0.94 | −2.14/0.99 | 1.00/0.77 | −0.61/0.99 |
| 18 | −0.6 | −3 | −5.2 | 89.3 | $3.40 \cdot 10^3$ | $5.25 \cdot 10^3$ | 0.97/0.98 | −0.894/0.80 | 1.00/0.99 | −0.19/0.90 | 1.01/0.91 | −51.3/0.94 |

geometric progression, whereas the $[f_{all}, f_{finish}]$ segment in test 10 is approximated using a power function.

The impact of the *ND* dimension of the search space is shown in Fig. 4 (tests 11–14). Expectedly, convergence slowed down with an increase in *ND*. Furthermore, for convex OF (tests 2, 11, and 12), a two-fold increase in *ND* leads to an equal upward shift in ERT on a logarithmic scale, while maintaining linear convergence. No clear patterns were observed for concave OF (tests 5, 13, and 14).

Introducing the loss of smoothness along the semi-infinite lines starting from the LE, in the case of a convex objective functions (tests 2, 15, and 16), results in a slowing down of the search process. This now reaches a limit of $f = 10^{-8}$ in not all tests. In the case of concave OFs (tests 5, 17, and 18), there is an increased minimum value, $f_{all}$, achieved at least once. In tests 15–18, the segment $[f_{start}, f_{all}]$ is approximated by a geometric progression with good accuracy. The power function contributes significantly to the approximation of the segments $[f_b, f_{finish}]$ and $[f_b, f_{finish}]$, as values for $q_2$ and $q_3$ are close to 1. However, uncommonly large values of $|p_2|$ and $|p_3|$ may occur in the results of some tests. This is due to gradual changes in ERT over the corresponding segments.

## CONCLUSIONS

Based on the results of the experiments, it can be inferred that dividing the range of ERT thresholds into segments and approximating the function $f$(ERT) using the proposed form (9) for each segment allows certain patterns to be identified. The presence of convexity in at least some directions ensures a minimum threshold of $f = 10^{-8}$ to be achieved. In the case of segments $[f_{start}, f_{all}]$ which include the values of the OFs achieved in all searches, the convergence can be considered to be linear with a fairly high degree of accuracy. In the case of those segments $[f_{all}, f_{finish}]$, where the OF values are not achieved in all searches, convergence follows a power law of the ERT$^p$ form.

Some results raise questions. In certain segments, the OF values are not achieved in all searches. This is understandable in cases where the OF is anisotropic, such as in tests 10 and 15–17. However, in tests 4–6 and 14, where the OF is isotropic, all starting points should yield the same search result. It is possible that the variation in the final search results in this case is due to random rounding errors in the QNA calculations.

These studies, of course, do not provide a complete understanding of the dependence of QNA convergence characteristics on OF landscape properties since within these studies, these properties vary separately and within limited ranges. In order to obtain a more comprehensive understanding, landscape properties need to be varied in various combinations over a wide range of parameter values. This would require a significant amount of experimentation and data analysis. The implementation of such an approach, presumably, could be achieved through the automation of experiment planning, data collection, and processing, including the use of artificial intelligence techniques.

Another area of interest is the expansion of the method to other types of algorithms aimed at finding an optimal solution, particularly population-based algorithms. This would allow for the comparison of convergence characteristics among different algorithms, essential for selecting the most effective algorithms for various problem types.

## REFERENCES

1. Polyak B.T. *Vvedenie v optimizatsiyu* (*Introduction into Optimization*). Moscow: Nauka; 1983, 384 p. (In Russ.).
2. Nocedal J., Wright S. *Numerical Optimization*: 2nd ed. Springer; 2006, 684 p.
3. Dem'yanov V.F., Vasil'ev L.V. *Nedifferentsiruemaya optimizatsiya* (*Non-differentiable Optimization*). Moscow: Nauka; 1981, 384 p. (In Russ.).
4. Vorontsova E.A., Hilderbrand R.F., Gasnikov A.V., Stonyakin F.S. *Vypuklaya optimizatsiya* (*Convex Optimization*). Moscow: MPTI; 2021, 364 p. (In Russ.). ISBN 978-5-7417-0776-0
5. Puchinin S.M., Korolkov E.R., Stonyakin F.S., Alkousa M.S., Vyguzov A.A. Subgradient methods with B.T. Polyak-type step for quasiconvex minimization problems with inequality constraints and analogs of the sharp minimum. *Komp'yuternye issledovaniya i modelirovanie = Computer Research and Modeling*. 2024;16(1):105–122 (in Russ.). https://doi.org/10.20537/2076-7633-2024-16-1-105-122
6. Bento G.C., Mordukhovich B.S., Mota T.S., Nesterov Yu. Convergence of Descent Methods under Kurdyka-Lojasiewicz Properties. *arXiv. Preprint*. 2024. http://arxiv.org/pdf/2407.00812v1
7. Grimmer B., Jia Zh. Goldstein Stationarity in Lipschitz Constrained Optimization. *Optim. Lett*. 2025419:225–235. https://doi.org/10.1007/s11590-024-02158-1
8. Lewis A.S., Overton M.L. Nonsmooth optimization via quasi-Newton methods. *Math. Program.* 2013;141:135–163. https://doi.org/10.1007/s10107-012-0514-2
9. Tor A.H. Comparative numerical results on HANSO (Hybrid Algorithm for Nonsmooth Optimization). *arXiv. Preprint*. 2020. http://arxiv.org/pdf/2009.01037v1

10. Varelas K., Dahito M.-A. Benchmarking Multivariate Solvers of SciPy on the Noiseless Testbed. In: *GECCO 2019 Companion – The Genetic and Evolutionary Computation Conference*, Jul. 2019, Prague, Czech Republic. 2019. P. 1946–1954. https://doi.org/10.1145/3319619.3326891
11. Hansen N., Finck S., Ros R., Auger A. *Real-Parameter Black-Box Optimization Benchmarking 2009*: *Noiseless Functions Definitions*. [Research Report] RR-6829. INRIA; 2009. Available from URL: https://hal.inria.fr/inria-00362633v2. Accessed August 26, 2025.
12. Smirnov A.V. Investigation of influence of objective function valley ratio on the determination error of its minimum coordinates. *Russian Technological Journal.* 2023;11(5):57−67. https://doi.org/10.32362/2500-316X-2023-11-6-57-67
13. Smirnov A.V. The method of estimation of objective functions landscapes convexity during extremum search. *Russian Technological Journal.* 2025;13(2):121–131. https://doi.org/10.32362/2500-316X-2025-13-2-121-131
14. Hansen N., Auger A., Ros R., Mersmann O., Tusar T., Brockhoff D. COCO: A Platform for Comparing Continuous Optimizers in a Black-Box Setting. *Optim. Meth. Software.* 2021;36(1):114–144. https://doi.org/10.1080/10556788.2020.1808977
15. Wang H., Vermetten D., Ye. F., Doerr C., Back T. IOHanalyzer: Detailed Performance Analyses for Iterative Optimization Heuristics. *ACM Transaction on Evolutionary Learning and Optimization.* 2022;2(1):1–29. https://doi.org/10.1145/3510426

## СПИСОК ЛИТЕРАТУРЫ

1. Поляк Б.Т. *Введение в оптимизацию*. М.: Наука; 1983, 384 с.
2. Nocedal J., Wright S. *Numerical Optimization*: 2nd ed. Springer; 2006, 684 p.
3. Демьянов В.Ф., Васильев Л.В. *Недифференцируемая оптимизация*. М.: Наука; 1981, 384 с.
4. Воронцова Е.А., Хильдебранд Р.Ф., Гасников А.В., Стонякин Ф.С. *Выпуклая оптимизация.* М.: МФТИ; 2021, 364 с. ISBN 978-5-7417-0776-0
5. Пучинин С.М., Корольков Е.Р., Стонякин Ф.С., Алкуса М.С., Выгузов А.А. Субградиентные методы с шагом типа Б.Т. Поляка для задач минимизации квазивыпуклых функций с ограничениями-неравенствами и аналогами острого минимума. *Компьютерные исследования и моделирование*. 2024;16(1):105–122. https://doi.org/10.20537/2076-7633-2024-16-1-105-122
6. Bento G.C., Mordukhovich B.S., Mota T.S., Nesterov Yu. Convergence of Descent Methods under Kurdyka-Lojasiewicz Properties. *arXiv. Preprint*. 2024. http://arxiv.org/pdf/2407.00812v1
7. Grimmer B., Jia Zh. Goldstein Stationarity in Lipschitz Constrained Optimization. *Optim. Lett*. 2025419:225–235. https://doi.org/10.1007/s11590-024-02158-1
8. Lewis A.S., Overton M.L. Nonsmooth optimization via quasi-Newton methods. *Math. Program.* 2013;141:135–163. https://doi.org/10.1007/s10107-012-0514-2
9. Tor A.H. Comparative numerical results on HANSO (Hybrid Algorithm for Nonsmooth Optimization). *arXiv. Preprint*. 2020. http://arxiv.org/pdf/2009.01037v1
10. Varelas K., Dahito M.-A. Benchmarking Multivariate Solvers of SciPy on the Noiseless Testbed. In: *GECCO 2019 Companion – The Genetic and Evolutionary Computation Conference*, Jul. 2019, Prague, Czech Republic. 2019. P. 1946–1954. https://doi.org/10.1145/3319619.3326891
11. Hansen N., Finck S., Ros R., Auger A. *Real-Parameter Black-Box Optimization Benchmarking 2009*: *Noiseless Functions Definitions*. [Research Report] RR-6829. INRIA; 2009. URL: https://hal.inria.fr/inria-00362633v2. Дата обращения 26.08.2025. / Accessed August 26, 2025.
12. Смирнов А.В. Исследование влияния степени овражности целевой функции на погрешность определения координат ее минимума. *Russian Technological Journal.* 2023;11(6):57−67. https://doi.org/10.32362/2500-316X-2023-11-6-57-67
13. Смирнов А.В. Метод оценки выпуклости рельефа целевых функций в процессе поиска экстремума. *Russian Technological Journal.* 2025;13(2):121–131. https://doi.org/10.32362/2500-316X-2025-13-2-121-131
14. Hansen N., Auger A., Ros R., Mersmann O., Tusar T., Brockhoff D. COCO: A Platform for Comparing Continuous Optimizers in a Black-Box Setting. *Optim. Meth. Software.* 2021;36(1):114–144. https://doi.org/10.1080/10556788.2020.1808977
15. Wang H., Vermetten D., Ye F., Doerr C., Back T. IOHanalyzer: Detailed Performance Analyses for Iterative Optimization Heuristics. *ACM Transaction on Evolutionary Learning and Optimization.* 2022;2(1):1–29. https://doi.org/10.1145/3510426

### About the Author

**Alexander V. Smirnov,** Cand. Sci. (Eng.), Professor, Department of Telecommunications, Institute of Radio Electronics and Informatics, MIREA – Russian Technological University (78, Vernadskogo pr., Moscow, 119454 Russia). E-mail: av_smirnov@mirea.ru. Scopus Author ID 56380930700, RSCI SPIN-code 1616-0120, https://orcid.org/0000-0002-2696-8592

### Об авторе

**Смирнов Александр Витальевич,** к.т.н., доцент, профессор кафедры телекоммуникаций, Институт радиоэлектроники и информатики, ФГБОУ ВО «МИРЭА – Российский технологический университет» (119454, Россия, Москва, пр-т Вернадского, д. 78). E-mail: av_smirnov@mirea.ru. Scopus Author ID 56380930700, SPIN-код РИНЦ 1616-0120, https://orcid.org/0000-0002-2696-8592

*Translated from Russian into English by Kirill V. Nazarov*
*Edited for English language and spelling by Dr. David Mossop*