

Mix-matrix Method in Problem of Discrete Optimization

Iakov Karandashev and Boris Kryzhanovsky

Center of Optical Neural Technologies

Scientific Research Institute for System Analysis, Russian Academy of Sciences

Moscow, Russia

Yakov.Karandashev@phystech.edu, kryzhanov@mail.ru

Abstract—The problem of a quadratic functional minimization in the configuration space of N binary states is considered. In order to increase the efficiency of the random-search algorithm, we suggest to vary the attraction area of the deepest minima of the functional by changing the matrix T it is based on. The new matrix M , called *mix-matrix*, is a mixture of T and T^2 . We demonstrate that this brings about changes of the energy surface: deep minima displace very slightly in the space (the Hamming distance of the shift is of about $0.01 \cdot N$), they become still deeper and their attraction areas grow significantly. The experiment shows that use of the approach results in a considerable displacement of the spectrum of sought-for minima to the area of greater depths, while the probability of finding the global minimum increases abruptly (by a factor of 10^3 in the case of a two-dimensional Ising model)

Keywords—quadratic optimization; binary optimization; combinatorial optimization; area of attraction; local search; random search; energy landscape transformation; mix-matrix

I. INTRODUCTION

The goal of this paper is to improve the efficiency of a random search procedure used to solve binary minimization problems. In this class of problems, the solution is reduced to the minimization of the quadratic functional $E(S)$ constructed from a given $N \times N$ matrix T in the N dimensional configuration space of states $S = (s_1, s_2, \dots, s_N)$ with discrete variables $s_i = \pm 1$, $i = 1, 2, \dots, N$. Many discrete programming problems, such as graph partitioning, graph coloring, traveling salesman problem etc., are reduced to this problem [1-2]. In addition, this problem arises in condensed matter physics where the search of the ground state is important for understanding of a disordered system structure [2].

It is well known that there is no polynomial algorithm for solving this problem, i.e., it is impossible to find a global minimum in polynomial time (the problem is NP -hard). Attempts are usually made to improve the efficiency of the random search procedure by modifying the dynamics of a descent over the landscape [1–3] described by $E(S)$. In contrast to this approach, we propose not to change the dynamics of landscape descent but rather to transform the energy landscape itself so as to increase the radius of the attraction domain of the global minimum (and of other minima comparable in depth with the global one).

In previous work [9], we consider the simplest transformation, namely, the raising of T to the power $k = 2, 3, \dots$. This approach was found to be productive: due to the landscape transformation, the spectrum of found minima is strongly shifted towards the deep side and the probability of finding the global minimum increases by 10^3 times. It was shown that the optimal value of power is $k = 3$. But the algorithm is unstable at $k \geq 3$: for the most part (about 70% of instances) the probability of finding global minima increases more than by 3 orders of magnitude in average, but sometimes (the rest 30%) it may decrease up to zero.

In present paper, we suggest to use a mix-matrix M , i.e., a mixture of T and T^2 . We claim that this yields a more reliable approach.

The efficiency of the algorithm proposed is rigorously substantiated only for “random” matrices, whose elements generated as independent random variables. The application of the algorithm to matrices of other types is heuristic.

The paper is constructed as follows. Section 2 includes the problem definition. Some preliminaries concerned the energy landscape of quadratic functional are given in Section 3. We describe the suggested idea regarding mix-matrix in Section 4. In Section 5, it is shown how the mix-matrix transforms the energy landscape of quadratic functional. Section 6 contains the obtained results for matrices of two types (uniform matrices and matrices of 2D Ising model).

II. PROBLEM DEFINITION AND MINIMIZATION PROCEDURE

The standard statement of the binary minimization problem is as follows. Given an $N \times N$ matrix T , find an N -dimensional configuration vector $S_m = (s_{m1}, s_{m2}, \dots, s_{mN})$, $s_{mi} = \pm 1$, $i = 1, 2, \dots, N$, that minimizes the energy functional $E(S)$:

$$E(S) = -\frac{1}{\sigma_T N^2} \sum_{i=1}^N \sum_{j=1}^N T_{ij} s_i s_j, \quad (1)$$

where σ_T is the standard deviation of the matrix elements T_{ij} . Functional (1) can be symmetrized. For this reason, without loss of generality, we assume that the matrix T_{ij} is symmetric and its diagonal elements are zero ($T_{ii} = 0$).

The minimization procedure is based on the Hopfield model [4], which is the core of most binary minimization algorithms. This is a one-dimensional system of N spins, whose interaction is defined by the energy functional $E(S)$. The standard (asynchronous) dynamics of the model can be described as follows (The full description is shown in Listing 1). The local field $h_i = -\partial E(S) / \partial s_i$ acting on the arbitrarily chosen i -th spin is calculated as

$$h_i = \frac{1}{\sigma_T N^2} \sum_{j \neq i}^N T_{ij} s_j \quad (2)$$

If $h_i s_i < 0$, the state of the spin is updated according to the decision rule $s_i = \text{sign } h_i$. This procedure is sequentially applied to all the neurons until the network converges to a stable state S_m . This dynamics is a descent over the energy landscape $E(S)$, which is a complete analogue of the coordinate-wise gradient descent in a real space.

Listing 1. The program code of the dynamics.

```

algorithm Asynchronous Neural Network Dynamics
    Initialize S = (s1, s2, ... sN) si = ±1
begin
    for i = 1 : N
        hi = ∑j≠i Tij sj           %calculate local fields
    end for
    flip = 1
    while (flip > 0)
        flip = 0
        for i = 1 : N
            if (hi si < 0) then
                si = -si           %reverse spin
                for j = 1 : N, j ≠ i
                    hj = hj + 2Tij si %refresh fields
                end for
                flip = flip + 1
            end if
        end for
    end while
end
    
```

NP -complete problems are known to have a huge number of local minima. In order to find a global one we have to use the random search. The random search procedure is described as follows. Given an arbitrarily initial state of the network, the nearest local minimum is found. This procedure is repeated until a minimum with an acceptable depth is found. The efficiency of the random search procedure is evaluated by the probability of finding the global minimum, by the rate of finding a minimum with a given depth, or by the mean depth of the minima found.

III. PRELIMINARIES

Before transforming the energy landscape, we establish the basic relations associated with the depth of the global (local) minimum, which underlie the subsequent argument.

The first relation is a constraint on the depth of the minimum. Let $S_0 = (s_{01}, s_{02}, \dots, s_{0N})$ be the configuration corresponding to the global minimum $E_0 = E(S_0)$. We

extract from T the term T_0 that is responsible for the formation of this minimum. To this end, T is represented as

$$T = T_0 + T_1, \quad T_0 = r_0 \sigma_T S_0^+ S_0 \quad (3)$$

The statistical weight r_0 is found from the condition that the elements of T_0 and T_1 do not correlate. Calculating the covariance of the matrix elements and setting it equal to zero, we obtain

$$r_0 = -\frac{E_0 + \bar{T} \delta}{1 - \delta^2}, \quad \delta = \frac{1}{N^2} \left[\left(\sum_{i=1}^N s_{0i} \right)^2 - N \right], \quad (4)$$

where \bar{T} is the mean of the elements of T and δ is a variable with a zero mean and a small standard deviation $\sigma_\delta = \sqrt{2}/N$. For simplicity, we set $\bar{T} = 0$ and $\delta = 0$ (the generalization to other cases is obvious). Then (4) yields the relation

$$E_0 = -r_0 \quad (5)$$

The variances of the elements of T_0 and T_1 are $\sigma_0^2 = r_0^2 \sigma_T^2$ and $\sigma_1^2 = \sigma_T^2 - \sigma_0^2$. Therefore, we have managed to present the random matrix T as the sum of two independent random matrices T_0 and T_1 . Moreover, (3) and (4) imply that $S_0 T_1 S_0^+ = 0$, which suggests that the contribution of T_1 to E_0 is strictly zero; i.e., the minimum in S_0 is caused only by the contribution of T_0 .

Following [5], we continue decomposition (3) and represent the matrix as a weighted sum of exterior products of random vectors:

$$T = \sigma_T \sum_0^\infty r_m S_m^+ S_m, \quad \sum_0^\infty r_m^2 = 1.$$

For this type of matrices, it was shown in [6] that any of the vectors S_m present in the decomposition of T is a minimizer of functional (1) if and only if its weight r_m is larger than the critical value

$$r_c = \frac{1}{2\sqrt{0.138N}} \quad (6)$$

This assertion is concerned primarily with the point S_0 , which by definition is a minimizer of functional (1) and satisfies the relations

$$1 \geq r_0 \geq r_c, \quad E_c \geq E_0 \geq -1, \quad E_c = -r_c$$

The second necessary relation obtained in [7] is that, as the depth of minimum E_0 increases, its width increases as well and, accordingly, the probability of finding this minimum grows as $P(E_0) \sim \exp(-NE_c^2 / E_0^2)$

As a result, we have established the following two relations:

- For a larger weight r_0 , the minimum E_0 is deeper and the probability of finding it is higher.
- S_0 can be a minimum only if $r_0 \geq r_c$; i.e., the depth of the minimum is larger than the critical value $|E_c|$.

These relations suggest the direction of improving the efficiency of the random search algorithm: the energy landscape (1) has to be transformed so as to increase the depth of the global minimum and, accordingly, to increase the probability of finding it.

IV. THE ALGORITHM

In this section we describe the proposed minimization algorithm. The main idea underlying the algorithm is the transformation of energy landscape of the functional. The surface described by the quadratic form $E(S)$ can be transformed only by transforming the underlying matrix.

Let us define the *mix-matrix* M as:

$$M = \frac{1-z}{\sigma_T} T + \frac{z}{\sigma_{2T}} T^2, \quad (7)$$

where T^2 is obtained by raising T to the second power and setting the diagonal elements equal to zero, σ_T and σ_{2T} are the standard deviations of matrices T and T^2 respectively. Substitute the new matrix into (1). Changing the parameter z from 0 to 1, we pass from the matrix T to T^2 . Accordingly, the landscape described by $E(S)$ is transformed into that described by:

$$E_z(S) = -\frac{1}{\sigma_M N^2} \sum_{i=1}^N \sum_{j \neq i}^N M_{ij} s_i s_j \quad (8)$$

where σ_M is the standard deviation of M_{ij} . Obviously, under the landscape transformation, the global minimum is shifted in space and its depth and the width of the attraction domain change as well.

Accordingly, we propose the following minimization algorithm. Firstly, we choose a value z , then construct the mix-matrix (7) and accordingly the functional $E_z(S)$. Then we start the minimization procedure consisting of two steps:

- At the first step, a descent over $E_z(S)$ is performed and a configuration S_{zm} is found that minimizes $E_z(S)$.
- The second step involves correction, namely, from the point S_{zm} , we descend over $E(S)$ to the nearest minimum S_m of $E(S)$.

Listing 2. The program code of the proposed algorithm.

```

algorithm Mix - matrix algorithm
    Initialize S = (s1, s2, ... sN) si = ±1
    Initialize the mix - matrix M with certain z
begin
    %1.Descent over transformed landscape
    for i = 1 : N
        hi = ∑j≠i Mijsj %calculate local fields
    end for
    flip = 1
    while (flip > 0)
        flip = 0
        for i = 1 : N
            if(hisi < 0) then
                si = -si %reverse spin
                for j = 1 : N, j ≠ i
                    hj = hj + 2Mijsi %refresh fields
                end for
                flip = flip + 1
            end if
        end for
    end while
    %2.Descent over initial landscape
    for i = 1 : N
        hi = ∑j≠i Tijsj %calculate local fields
    end for
    flip = 1
    while (flip > 0)
        flip = 0
        for i = 1 : N
            if(hisi < 0) then
                si = -si %reverse spin
                for j = 1 : N, j ≠ i
                    hj = hj + 2Tijsi %refresh fields
                end for
                flip = flip + 1
            end if
        end for
    end while
end
    
```

The descent over $E_z(S)$ is performed as described above: we calculate the local field of the i^{th} spin $h_i^{(z)} = -\partial E_z(S) / \partial s_i$ and, if $h_i^{(z)} s_i < 0$, the state of the spin is updated according to the decision rule $s_i = \text{sign } h_i^{(z)}$. The full description of the algorithm is given in Listing 2.

In previous work [9] we consider the simplest transformation, namely, when $M = T^k$, $k = 2, 3, 4, 5$. It was shown that the optimal value of power is $k = 3$. In this case the probability of finding global minima increases by 3 orders of magnitude for the most part (about 70% of instances). But sometimes (the rest 30% ones) it may decrease up to zero due to vanishing a minimum near S_0 .

As a result of this, in present paper, we introduce a mix-matrix (7), i.e., a mixture of T and T^2 , and vary the parameter z from 0 to 1. This yields a more reliable approach.

We will show that at $z \approx 0.5$ the proposed transformation leads to significant increase of the global minimum depth, while the shift from the minimum is smaller (1-2% of N) than in case $M = T^3$ (3% of N).

V. CORRECTNESS OF THE ALGORITHM

The algorithm is substantiated only for “random” matrices, whose elements are independent random variables. The application of the algorithm to matrices of other types is heuristic.

A. The deepening of the minima.

Let us show that the landscape transformation leads to a deeper minimum. Consider the energy $E_{z_0} = E_z(S_0)$ at the point S_0 . Following (3), the mix-matrix M is represented as

$$M = (1-z) \frac{T_0 + T_1}{\sigma_T} + z \frac{T_0^2 + T_0 T_1 + T_1 T_0 + T_1^2}{\sigma_{2T}}$$

In view of $S_0 T_1 S_0^+ = 0$ and $\sigma_M^2 = (1-z)^2 + z^2$, we then derive from (8) that

$$E_{z_0} = \bar{E}_{z_0} + R \quad (9)$$

where

$$\bar{E}_{z_0} = - \frac{(1-z)r_0 + zr_0^2 \sqrt{N}}{\sqrt{(1-z)^2 + z^2}}$$

$$R = \frac{1}{N^2 \sqrt{(1-z)^2 + z^2}} \sum_{i=1}^N \sum_{j \neq i}^N \left(\frac{(1-z)T_1}{\sigma_T} + \frac{zT_1^2}{\sigma_{2T}} \right)_{ij} s_{0i} s_{0j}$$

In the limit of $N \gg 1$, E_{z_0} can be viewed as a normally distributed quantity with the mean value \bar{E}_{z_0} and the relatively small noise R of standard deviation $\sigma_R = 1/N$. The ratio:

$$\frac{\bar{E}_{z_0}}{E_0} = \frac{(1-z) + z\sqrt{N}r_0}{\sqrt{(1-z)^2 + z^2}} \quad (10)$$

shows how many times the average value of the modified functional at point S_0 more than the initial functional value at the same point. Taking into account $\sqrt{N}r_0 \approx 1.35$, it is obvious that at any value of z expression (10) is larger than unit, hence when $N \gg 1$ one can be sure that the minimum becomes deeper. Fig. 1 confirms this. The largest deepening ($E_{z_0} \approx 1.6E_0$) is observed at $z \approx 0.6$.

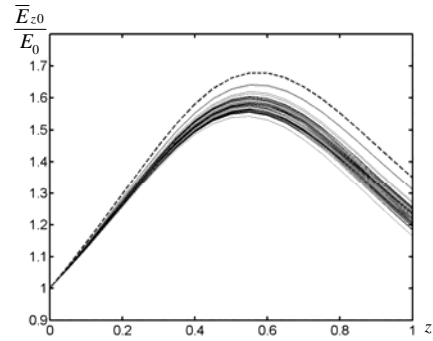


Figure 1. The decrease of energy in the point S_0 (global minimum) due to energy landscape transformation (mix-matrix with T^2). The dashed line is theoretical (10). Other lines are experimental for 50 random instances with uniform matrices.

B. The shift of the minima.

Let us estimate the shift of the minimum under the landscape transformation. The mean shift can be represented as

$$d = N \cdot P,$$

where $P = \Pr\{s_{0i} h_i^{(z)} < 0 \mid s_{0i} h_i > 0\}$ is the probability that the directions of the spin s_{0i} and the local field $h_i^{(z)}$ do not coincide. Omitting the unnecessary constants, the value $s_{0i} h_i^{(z)}$ can be represented as

$$h_i^{(z)} s_{i0} = (1-z)Nr_0 + zN^{3/2}r_0^2 + H \quad (11)$$

where

$$H = \sum_{i=1}^N \left(\frac{(1-z)T_1}{\sigma_T} + \frac{z(Nr_0\sigma_T T_1 + T_1^2)}{\sigma_{2T}} \right)_{ij} s_{0i} s_{0j}$$

In view of (11), P is expressed in terms of the error function:

$$P = \frac{1}{2\Phi(\gamma)\sqrt{2\pi}} \int_0^\infty dx e^{-\frac{1}{2}(x-\gamma)^2} (1 - \Phi(\alpha\sigma x)), \quad (12)$$

where $\Phi(\cdot)$ is the probability integral and

$$\gamma = \sqrt{N}r_0 / \sigma \approx 1.9,$$

$$\sigma = \sqrt{\frac{1}{\sigma_T^2 N} \sum_{i=1}^N h_i^2} \approx 0.7,$$

$$\alpha = \frac{1-z}{z} + \sqrt{N}r_0$$

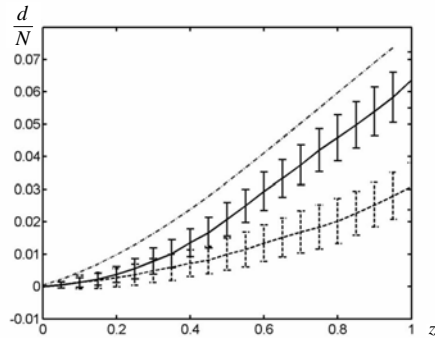


Figure 2. The shift (in bits) of the global minimum as a function of z (mix with T^2). The curves with error bars were obtained by experiment for two types of matrices: matrix with uniformly distributed elements (solid line) and 2D Ising matrices (dashed lines). The dash-dot line is theoretical (12).

Note that at $z=0$ the functional $E_{z=0}(S)$ coincides with initial $E(S)$ and therefore the shift is absent, this agrees ($d = N \cdot P = 0$) with (12).

The formula (12) describes a monotone increase of the minimum shift with growing z in view of enlarging functional transformation. This corresponds to a common sense and is proved by experiment (see fig. 2).

Expressions (9)–(12) suggest the following conclusions. With a high probability, the landscape transformation leads to deeper minima and, as a result, to a higher probability of finding them. Moreover, the depth increase (10) is larger for a larger initial depth $|E_0| \approx r_0$. In other words, deep minima become even deeper and the probability of finding them increases, while shallow minima become shallower (or disappear at all) and the probability of finding them is reduced. This means that the spectrum of minima found by the algorithm shifts considerably toward the global minimum, and the probability of finding the latter increases considerably. The spatial minima displacements caused by the transformation are relatively small: it follows from (12) that the smallest shifts are expected for the deepest minima.

VI. RESULTS

The efficiency of the two-step descent algorithm was verified for z ranging from 0 to 1 for matrices of size $N = 100, \dots, 500$ of two types:

- matrices with random elements uniformly distributed within $(-1; 1)$;
- matrices of 2-dimensional Ising model with [2].

During numerical experiments we built a mix-matrix for different values of z from 0 to 1 equally spaced with $\Delta z = 0.05$. The results were averaged over 50 random instances of each size and type.

The computational complexity of the algorithm is $O(N^2)$. In experiments, we used the same algorithm realization both for sparse and dense matrices, although it is possible to reduce the complexity up to $O(N)$ in case of 2D Ising matrices.

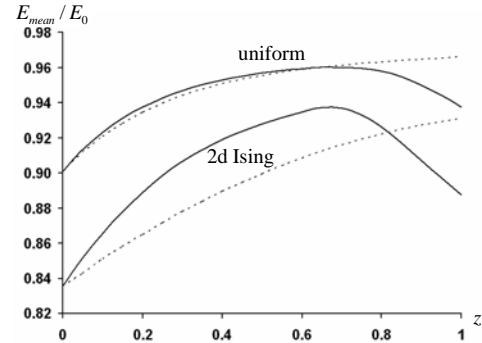


Figure 3. The mean value E_{mean} of energy of local minima found with the proposed two-step algorithm. The solid lines are for mix-matrices with T^2 . The dashed lines are for mix-matrices with T^3 . The curves are drawn for two types of matrices: uniform matrices (on top) and 2D Ising matrices. The value of E_{mean} is divided by the energy of global minimum E_0 and does not depend on the problem dimension N .

In addition, some part of the running time is spent on matrix multiplication. Nevertheless, each our experiment took no more than one hour for $N \approx 500$.

Each experiment included $N_{runs} = 10^6$ runs. Each run resulted in a local minimum. We chose two parameters to trace: the mean energy E_{mean} of the minima found and the probability of finding a minimum in the interval of energy close to global one $E \in [-1; -0.99]$, where -1 corresponds to E_0 .

In experiments, we try to use not only the mix-matrix (7) but also mix with T^3 . In this case the mix-matrix was constructed in the same manner (7) but T^2 was replaced with T^3 .

The numerical results are shown in Figs. 3 and 4.

Fig. 3 demonstrates how the mean value of energy E_{mean} of minima found for different z changes. It is interesting that the value E_{mean}/E_0 does not depend on the problem dimension but the type of the matrix.

As we can see from Fig. 3, E_{mean} comes near E_0 with increasing z . We observe the maximum of E_{mean}/E_0 at $z \approx 0.7$ for mix-matrix with T^2 and the monotone growth of E_{mean}/E_0 for mix-matrix with T^3 up to $z = 1$.

Fig. 4 shows how many times the probability of finding minima with energy differed from the global one less than 1% increases. For demonstration purpose, we chose the maximal possible problem dimensions, which we can cope with. For 2D Ising matrices the probability of finding minima of energy $E \in [-1; -0.99]$ is not greater than $P_1 = 3 \cdot 10^{-7}$ for $N = 12 \times 12$. For uniform matrices the maximal dimension is $N = 500$ (the probability $P_1 = 3 \cdot 10^{-5}$). The probability obtained with the proposed algorithm was denoted by P_{new} . As we can see from fig. 4,

the difference between P_{new} and P_1 turned out to be enormous – approximately 3 orders of magnitude.

An interesting fact is that for uniform matrices the T^2 and T^3 curves almost coincide (see. fig. 3-4), and they start to diverge when $z > 0.7$ only. For Ising matrices, we have another picture: mix with T^2 prevails over mix with T^3 up to $z \approx 0.8$ and after that vice versa.

It can be also seen from Fig. 4 that with increasing z the dispersion rises, and this can lead to the instability of the algorithm, i.e., the transformation may change the search procedure for the worse in some cases.

VII. CONCLUSION

Finally, we formulate the minimization algorithm proposed.

The preliminary phase consists of the following steps. The original matrix T is symmetrized (if it is initially not symmetric) and its diagonal elements are set to zero. The matrix is raised to the k^{th} power ($k = 2$ or 3) and the diagonal elements in the resulting matrix T^k are set to zero. Afterwards the matrices T and T^k are normalized on unit dispersion and mixed in accordance with (7), and the mix-matrix M is obtained. It depends on the chosen parameter $z \in (0,1)$. The functional $E_z(S)$ is constructed from M according to (8).

After the preliminary phase, the random search procedure based on the two-step descent algorithm is executed. Specifically, at the first step, a descent over surface $E_z(S)$ is performed from a random initial configuration to the nearest local minimum S_{zm} of $E_z(S)$. The second step involves correction: from the point we descend over surface $E(S)$ to the nearest local minimum S_m of $E(S)$, which is, as a rule, located near S_{zm} .

The simplest Hopfield neural network dynamics [4] was chosen as a descent dynamics (nevertheless, it can be arbitrary).

A comparison shows that the efficiency of the minimization algorithm is improved substantially due to the landscape transformation.

It was shown that we succeeded in decreasing the value $(E_0 - E_{mean})/E_0$ (difference between the mean energy of found minima and global one) by half when $k = 2$ and $z = 0.7$ (see fig. 3).

Due to the proposed method the probability of finding suboptimal solutions with energy differed from the optimum less than 1% increases by 2.5 orders of magnitude for

uniform (full) matrices of dimension $N = 500$ and by more than 3 orders for (sparse) matrices of Ising model of dimension $N = 12 \times 12$.

Finally, it seems to be attractive to use the proposed algorithm as a preliminary stage of any sort of genetic algorithms. Indeed, one can run our algorithm with different values of z to obtain a set of minima. Most of the minima must be deep and some of them may lie near the global minimum. Therefore, they are the good candidates for being parents.

ACKNOWLEDGMENT

The work was supported by the program of the Presidium of the Russian Academy of Sciences (project 2.15) and in part by the Russian Basic Research Foundation (grant 12-07-00295).

REFERENCES

- [1] B.W. Kernighan and S. Lin, "An Efficient Heuristic Procedure for Partitioning Graphs", Bell System Tech. Journal, 49, pp. 291–307, 1970.
- [2] A.K. Hartmann and H. Rieger, New Optimization Algorithms in Physics (Wiley, Weinheim), 2004.
- [3] C. Dang, W. Ma, and J. Liang, "A deterministic annealing algorithm for approximating a solution of the min-bisection problem", Neural Networks 22 (1), pp. 58–66, 2009.
- [4] J.J. Hopfield, "Neural Networks and physical systems with emergent collective computational abilities", Proc.Nat.Acad.Sci.USA. vol. 79, pp. 2554-2558, 1982.
- [5] B.V. Kryzhanovsky, "Expansion of a matrix in terms of external products of configuration vectors", Optical Memory and Neural Networks 6 (4), pp. 187–199, 2007.
- [6] D.J. Amit, H. Gutfreund, and H. Sompolinsky, "Spin-glass models of neural networks", Phys. Rev. A, vol. 32, pp. 1007-1018, 1985; Annals of Physics, vol. 173, pp. 30-67, 1987
- [7] B.V. Kryzhanovskii, B.M. Magomedov, and A.L. Mikaelyan, "A Relation Between the Depth of a Local Minimum and the Probability of Its Detection in the Generalized Hopfield Model", Doklady Mathematics, vol. 72, N3, pp. 986-990, 2005.
- [8] B.V. Kryzhanovsky and V.M. Kryzhanovsky, "The shape of a local minimum and the probability of its detection in random search", Lecture Notes in Electrical Engineering, Vol. 24, pp. 51-61, 2009.
- [9] Y.M. Karandashev and B.V. Kryzhanovsky, "Transformation of Energy Landscape in the Problem of Binary Minimization", Doklady Mathematics, v. 80, No. 3, pp. 927-931, 2009.

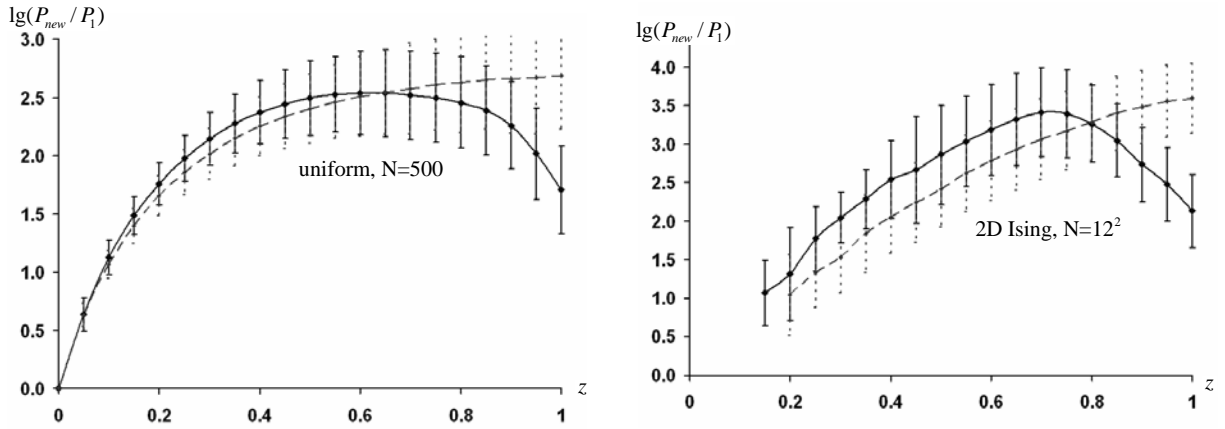


Figure 4. The common logarithm of the ratio of probabilities of hitting the energy interval $E \in [-1, -0.99]$. The solid lines are for mix-matrices with T^2 . The dashed lines are for mix-matrices with T^3 . In the left panel the results for uniform matrices of $N = 500$ ($P_1 \approx 3 \cdot 10^{-5}$). In the right panel the results for 2D Ising matrices of $N = 144$ ($P_1 \approx 2.6 \cdot 10^{-7}$). Note, that when z is too small, the algorithm does not find the global minimum in some instances, so the points are missed.