Study of the point scattering uniform algorithms in R^{40} space

A A Stupina[1,2,3]**, A A Pavlenko**[1]**, D V Tikhonenko**[1]**, V V Tynchenko**[1,2]**, V V Bukhtoyarov**[1,2] **and S V Tynchenko**[1,2]

[1]Reshetnev Siberian State University of Science and Technology, 31, KrasnoyarskyRabochy Av., Krasnoyarsk, 660037, Russian Federation
[2] Siberian Federal University, 79, Svobodny pr., Krasnoyarsk, 660041, Russian Federation
[3] Krasnoyarsk State Agrarian University, 90, Mira pr., 660049, Krasnoyarsk, Russia

E-mail: ms_f1_pavlenko@mail.ru

Abstract. The use of randomness in the spread of points in the R^{40} space gives doubts about the stability of using these spreads and the stability of global optimization algorithms predictions that are based on these spreads. Uniformity of the following initial points scatter algorithms is analyzed: LPτ sequence, UDC sequence, uniform random spread in R^{40}. The uniformity of the spread was determined by the distance of the points from the centers of the grid cells in two-dimensional coordinate planes of the R^{40} cube space and by the uniformity of the projections of the points on the coordinate axes in these planes. We identified the features of using the points spread algorithms when the number of points is multiple to two and not multiple to two. The UDC sequence is the best initial point spread algorithm in the R^{40} space by two uniformity factors. LPτ sequence is at the second place and recently used uniform random scatter is at the third place.

1. Introduction

It is a very nontrivial task to determine the uniformity of the spread in the n-dimensional space [1], where n > 3. Therefore, we will evaluate the definition of the inconsistency of the spreads to the uniform distribution law in n-dimensional space in two ways.

First, we will consider the two-dimensional coordinate planes of the cube in the R40 space and estimate the distance from the centers of the grid cells. Secondly, we will split each n-dimensional coordinate axis into N intervals, where N is the total number of scattered points and look at the number of points that are projected onto the given coordinate axis in each such interval, and then evaluate the uniformity of the distribution of points in one-dimensional space.

The dimension is taken 40, as the most complex option for analysis [2]. We studied the number of points 8, 10, 12, 16, 20, 32, 40, 64, 80, 100 [2]. These points were studied in six planes, (1-2), (2-3), (3-4), (20-21), (25-26), (35-36) [2]. Were investigated two-dimensional coordinate planes of the cube in the space R40, when the number of points is a multiple of two and not a multiple of two [2]. For greater clarity, perpendiculars on coordinate axes are omitted from most of the scattered points on most graphs. On the right side of each graph, the average error of the uniformity of the projection of points on the coordinate axes (ERmid) is given, taking into account all 40 coordinate axes.

LPτ sequences are an algorithm for scattering points based on a matrix of irreducible Marshal polynomials. UDC sequence is an algorithm for the absolutely uniform distribution of points across all coordinates in a multidimensional space, regardless of the number of scattered points. Uniform random spread is a stochastic point spread algorithm using the normal distribution law [3].

2. Explanatory part

Figure 1 shows a block diagram of the LPτ sequence [4].

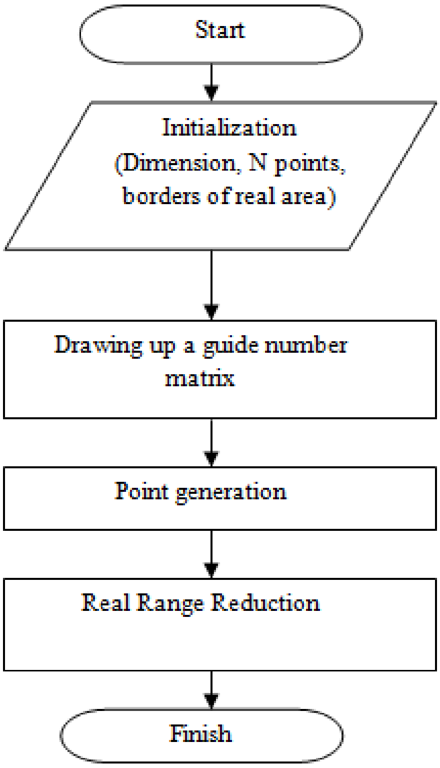

Figure 1. Block diagram of the generation algorithm LPτ– Sequences.

Recent research in this area was conducted by A.I. Diveev in 2018 [5]. His work was unsuccessful because he used the matrix of irreducible Marshal polynomials incorrectly. N is the number of scatter points in the R^{40} cube.

3. Experimental part

Figures 2 and 3 show the LPτ sequence plots in the two-dimensional coordinate plane 20-21.

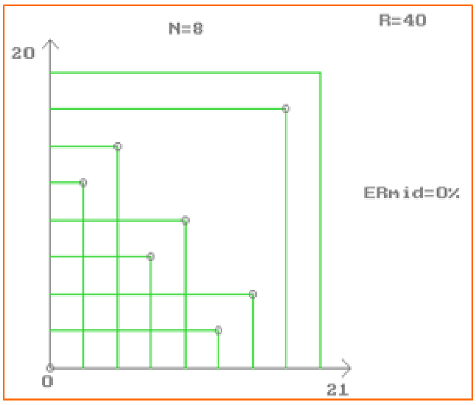

Figure 2. LPτ sequence. Plane (20-21) of the R^{40} cube at 8 points.

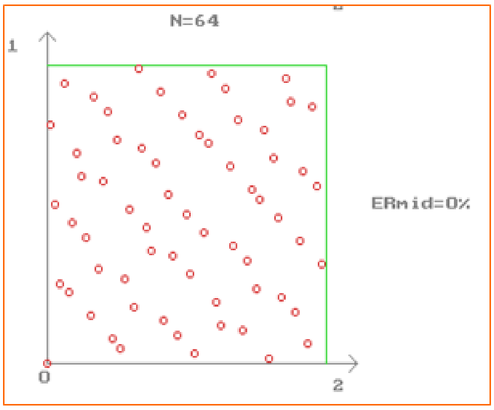

Figure 3. LPτ sequence. The plane (20-21) of the R^{40} cube with 64 points.

Figure 4 shows a graph of a uniform random scatter [6] in a two-dimensional coordinate plane 20-21. Figure 5 presents a graph of the UDC sequence [7]. Figure 3 shows the projections of points on the coordinate axes for the example of the LPτ sequence. Figures 3, 4, 5 present points for determining the uniformity of the spread of points by the equidistance of points from each other in a two-dimensional coordinate plane 20-21.

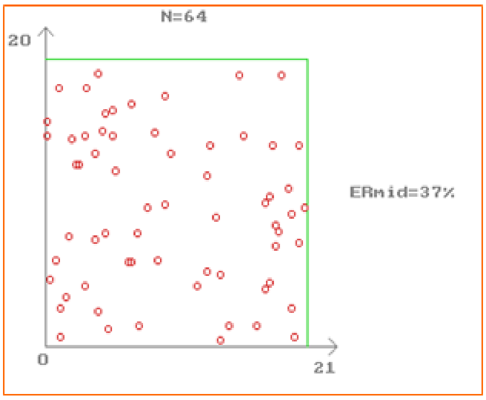

Figure 4. Uniform random scatter. The plane (20-21) of theR^{40} cube with 64 points.

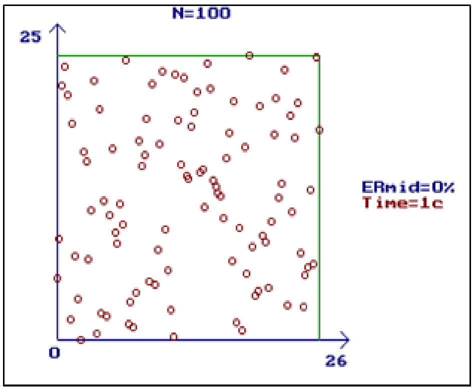

Figure 5. UDC spread. The plane (20-21) of the R^{40} cube at 100 points.

Table 1 shows the comparison of the uniformity average error of the spread of the points projections on the coordinate axes.

Table 1. Comparison of the average error of the uniformity of the spread of the projections of points on the coordinate axes.

The number of points, N	Dimension number	ER – rnd, %	ER – UDC, %	ER – LPτ, %	ER – rnd – middle, %	ER – LPτ – middle, %
8	1	25	0	0	25	0
	2	38	0	0		
	3	12	0	0		
	4	25	0	0		
	20	25	0	0		
	21	50	0	0		
	35	25	0	0		
	36	12	0	0		
100	1	41	0	20	36	20
	2	41	0	21		
	3	35	0	20		
	4	38	0	20		
	20	32	0	21		
	21	32	0	20		
	35	39	0	20		
	36	32	0	20		

The first column of Table 1 lists the number of scatter points in the R^{40} cube. The second column shows the number of dimensions that participated in the formation of the plane. The third one presents the differences (errors) of the projections of the scattering points along a random spread with a uniform distribution law in the R^{40} cube from the uniform distribution law along each coordinate axis involved in the formation of the plane. These errors are expressed as the variable ER-rnd and are determined as a percentage of the total number of projected points on a given coordinate axis.

The fourth column lists the same differences as in the third column, only when using LPτ sequences, expressed in the variable ER - LPτ. The fifth column lists the same differences as in the fourth column, only when using UDC sequences expressed in the variable ER - UDC. The sixth column shows the average error of the points projection spread uniformity distributed by a random mechanism with a uniform distribution law on the coordinate axes expressed in the variable ER - rnd - middle. The seventh column lists the average uniformity errors of the projection of LPτ sequence points on the coordinate axes expressed in the variable ER - LPτ- middle.

4. Results

The average error in the points projection spread uniformity in the LPτ sequence along each coordinate axis for $N=2^n$ equals 0%, and for $N \neq 2^n$ it's between 0% and 19%, with an error spread from 0% to 20%. The average error of the points projection spread uniformity in a random variation with a uniform distribution law on each coordinate axis, both for $N=2^n$ and $N \neq 2^n$ is between 33% to

36%, for an error spread from 20% to 62%. The average error in the points projection spread uniformity in the UDC scatter is always zero due to its special construction.

The error of the points projection spread uniformity with random variation and a uniform distribution on each coordinate axis with the same number of scattering points and different planes can vary from 12% to 50%, regardless of the number of scattering points, while the error of scattering the uniformity of points projections when using LPτ– sequences with one number of scattered points and different planes is 0% if $N=2^n$ and can vary from 0% to 19% if $N \neq 2^n$.

It can also be concluded that the greater the number of scattered points when using the LPτ sequence differs from $N=2^n$, the greater the error in the points projection spread along each coordinate axis.

5. Conclusion

Studies of the dispersion points on the factors of uniformity showed that LPτ sequence is behind the UDC sequence in terms of the distance of the points from the centers of the grid cells. According to the average error of the points projection spread uniformity on the coordinate axes, the places were distributed as follows: when $N=2^n$ the LPτ sequence and UDC sequence were the best, and with $N \neq 2^n$ the UDC sequence was better than the LPτ sequence, and the latter uniform random scatter showed itself from the worst side.

References

[1] Zaloga A N, Yakimov I S and Dubinin P S 2018 Multipopulation genetic algorithm for determining crystal structures using powder diffraction data *Journal of surface investigation: X-ray, synchrotron and neutron techniques* **1(12)** 128-34

[2] Stanovov V, Akhmedova S and Semenkin E 2018 Selective pressure strategy in differential evolution: exploitation improvement in solving global optimization problems *Swarm and Evolutionary Computation* 183-9

[3] Akhmedova S, Stanovov V and Semenkin E 2018 Soft island model for population-based optimization algorithms *International Conference on Swarm Intelligence* 68-77

[4] Zaloga A 2015 Genetic algorithm for automated X-Ray diffraction full-profile analysis of electrolyte composition on aluminium smelters *Informatics in control, automation and robotics 12th International Conference, ICINCO* 79-93

[5] Diveev A I and Konstantinov S V 2018 The problem of optimal control and its solution by the evolutionary algorithm of the "gray wolf" *Bulletin of the Russian University of friendship of peoples. Series: Engineering research* **1(19)** 59-67

[6] Yakimov I 2018 Application of evolutionary rietveld method based XRD phase analysis and a self-configuring genetic algorithm to the inspection of electrolyte composition in aluminum electrolysis baths *Crystals* **11(8)** 402

[7] Sharifpour E 2018 Zinc oxide nanorod loaded activated carbon for ultrasound assisted adsorption of safranin o: central composite design and genetic algorithm optimization *Applied organometallic chemistry* **2(32)** 4099