**International Journal of Educational Research and Information Science**
2015; 2(1): 1-6
Published online January 30, 2015 (http://www.openscienceonline.com/journal/eris)

# Monte Carlo method and its parallel computing technique in molecular gas dynamics

Yuri Ivanovich Khlopkov, Zay Yar Myo Myint, Anton Yurievich Khlopkov

Department of Aeromechanics and Flight Engineering, Moscow Institute of Physics and Technology, Zhukovsky, Russia

## Email address

khlopkov@falt.ru (Y. I. Khlopkov), zayyarmyomyint@gmail.com (Z. Y. M. Myint)

## To cite this article

Yuri Ivanovich Khlopkov, Zay Yar Myo Myint, Anton Yurievich Khlopkov. Monte Carlo Method and Its Parallel Computing Technique in Molecular Gas Dynamics. *International Journal of Educational Research and Information Science*. Vol. 2, No. 1, 2015, pp. 1-6.

## Abstract

The major function of parallel computing system is how to coordinate communication between the various processors. Some parallel computing techniques require specialized programming to permit the processors to work together in parallel. It can be seen that on Monte Carlo simulations, algorithms proceed by averaging large numbers of computed values. It is sometimes straightforward to have different processors compute different values, and then use an appropriate average of these values to produce a final result. In this paper give an introduction to the generalized Monte Carlo method and its related algorithms. This paper aims to describe Monte Carlo methods with parallel computing techniques and to present some Monte Carlo algorithms to solve some problems in gas dynamics.

## Keywords

Parallel Computing System, Monte Carlo Method, Computational Fluid Dynamics, Turbulent Flow, High-Speed Aircraft Technology

## 1. Introduction

We have already known that in the computer science where the subject of parallel computing is very prominent and includes such as parallel randomized algorithms and parallel simulation. Nowadays, as computer processors become cheaper and more plentiful, there is great potential for having them compute together in a coordinated application. A major point of parallel computing is how to coordinate communication between the various processors; indeed, some parallel computing techniques require specialized programming to permit the processors to work together in parallel. It can be seen that on Monte Carlo simulations, algorithms proceed by averaging large numbers of computed values. It is sometimes straightforward to have different processors compute different values, and then use an appropriate average of these values to produce a final answer.

The complex nonlinear structure of the collision integral and the large number of variables (seven in the general case) present severe difficulties for the analysis including the numerical analysis. The high dimension, the probabilistic nature of the kinetic processes, and complex molecular interaction models are the natural prerequisites for the application of the Monte Carlo methods. The first paper devoted to the Monte Carlo method was published as early as in 1873 [1]. It described the experimental determination of $\pi$ by a realization of the stochastic process of tossing a needle on a sheet of ruled paper. A striking example is the use of von Neumann's idea to simulate the neutron trajectories in the Los Alamos laboratory in 1940. Although the Monte Carlo methods require a large amount of computations, the absence of computers at that time did not discourage the researchers. The name of these methods comes from the capital of the Principality of Monaco, which is famous for its Casino; indeed, the roulettes used in the casino are perfect tools for generating random numbers. The first paper [2] that systematically expanded this method was published in 1949. In that paper, the Monte Carlo method was used to solve linear integral equations. It could easily be guessed that these equations were related to the problem of the passage of neutrons through

matter. In Russia, studies concerning the Monte Carlo methods appeared after the Geneva International Conference on the Peaceful Uses of Atomic Energy. One of the first Russian studies is [3].

The numerical statistical methods in rarefied gas dynamics developed in three directions: the use of the Monte Carlo methods to evaluate the collision integrals in the regular finite difference schemes for solving the kinetic equations; the direct statistical simulation of physical phenomena, which is subdivided into two approaches: the simulation of trajectories of test particles by the Haviland method [4] and the simulation of the evolution of the ensemble of particles by the Bird method [5]; the construction of a stochastic process using the Ulam–Neumann procedure [6] corresponding to the solution of the kinetic equation.

The revelation of the direct simulation of method Monte Carlo (DSMC) in various areas of the applied mathematics is connected, as a rule, with the necessity of solution of the qualitatively new problems, arising from the needs of practice. Such a situation appeared by the creation of the atomic weapon, at the initial stage of a mastering of space, by the investigation of the phenomena of atmospheric optics, of the physical chemistry, and of the modeling of turbulence (G. von Neumann, Metropolis N., Unlam S., Vladimirov V.S., Sobol I.M., Marchuk G.I., Ermakov, S.M., Mikhailov G.A., Bird G.A., Haviland J.K., Lavin M.D., Pullin D.I., Kogan M.N., Perepukhov V.A., Beloserkovskii O.M., Yanitskii V.E., Khlopkov Yu.I., Ivanov M.S. and Eropheev A.I.).

The aim of this paper is to present the Monte Carlo methods with parallel techniques. In this paper present some Monte Carlo algorithms to solve some parallel computation problems.

## 2. General Principle of Monte Carlo Methods

The general scheme of the Monte Carlo method is based on the central limit theorem, which states that, the random quantity

$$Y = \sum_{i=1}^{N} X_i$$

equal to the sum of a large number of random variables with the same expectation $m$ and the same dispersions $\sigma^2$. They have the normal distribution with the expectation $N$ and the variance $N\,\sigma^2$. Assume that we want to solve an equation or find the result of some process $I$. If we can construct the random variable $\xi$ with the probability density $p(x)$ such that the expectation of this variable is equal to the unknown solution $M(\xi) = I$, then we obtain a simple method for estimating the solution and its error:

$$I = M(\xi) \approx \frac{1}{N}\sum_{i=1}^{N} \xi_i \pm \frac{3\sigma}{\sqrt{N}}$$

This implies the general properties of the Monte Carlo methods as below:

- the absolute convergence to the solution with the rate $1/N$.
- an unfavorable dependence of the error $\varepsilon$ on the number of trials: $\varepsilon \approx 1/\sqrt{N}$ (to reduce the error by an order of magnitude, the number of trials must by increased by two orders of magnitude).
- the main method of reducing the error is the variance reduction; in other words, this is a good choice of the probability density $p(x)$ of the random variable $\xi$ in accordance with the physical and mathematical formulation of the problem.
- the error is independent of the dimensionality of the problem.
- a simple structure of the computation algorithm (the computations needed to realize a proper random variable are repeated $N$ times).

The structure of the random variable $\xi$ can be generally based on a physical model of the process that does not require a formulation of the controlling equations as in regular methods; this fact is increasingly important for modern problems.

We illustrate the main features of the Monte Carlo methods and the conditions under which these methods outperform the conventional finite difference methods or are inferior to them using the following example. Suppose that we want to evaluate the definite integral of a continuous function over the interval $[a, b]$:

To evaluate this integral using the Monte Carlo method, we construct a random variable with the probability density $p(x)$ such that

$$M(\xi) = \int_{-\infty}^{\infty} \xi \cdot p(x)dx$$

is equal to $I$. Now, if we set $\xi = f(x)/p(x)$ within the integration limits, then we have, by the central limit theorem,

$$I = \frac{1}{N}\sum_{i=1}^{N} \xi_i \pm \frac{3\sigma}{\sqrt{N}}$$

On the one hand, the evaluation of $I$ by above formula can be interpreted as the solution of mathematically stated problem; on the other hand, it can be interpreted as a direct simulation of the area under the function of $f(x)$. The evaluation of the one-dimensional integral $I_1$ by the Monte Carlo method corresponds to the computation of $I$ using the rectangular rule with the step $\Delta x \approx 1/\sqrt{N}$ and an error $O(\Delta x)$. If $f(x)$ is sufficiently "good", the integral $I_1$ in the one-dimensional case can be calculated accurate to $O(\Delta x^2)$ using the trapezoid rule, accurate to $O(\Delta x^3)$ using the parabolic rule, and to any desired accuracy without a considerable increase in the computational effort. In the multidimensional case, the difficulties in using schemes of a high order of accuracy increase; for this reason, they are rarely used for the calculation of $n$-dimensional integrals $I_n$ for $n \geq 3$.

Let us compare the efficiency of the regular and statistical

methods for the problem described above. Let $n$ be the dimensionality of the problem, $Y$ be the number of nodes on an axis, $R = Y^n$ be the total number of nodes for the regular methods, $q$ be the order of accuracy, $N$ be the number of statistical trials, and ν be the number of operations needed to process one node (to perform one statistical trial). Then, $\varepsilon_L = Y^{-q}$ is the error of the regular methods, $\varepsilon_K = N^{-1/2}$ is the error of the statistical methods, $L(\varepsilon) = \nu \cdot R = \nu \cdot \varepsilon^{n/q}$ is the number of operations when the problem is solved by a regular method, and $K(\varepsilon) = \nu \cdot N = \nu \cdot \varepsilon^{-2}$ is the number of operations when the problem is solved by the Monte Carlo method. Then, in the case of an equal number of operations needed to obtain a solution with the same accuracy using each of the methods, we have $n = 2q$. Therefore, for $n \geq 3$ and $q = 1$ (first-order schemes), the Monte Carlo methods are preferable. For other classes of problems, the relation between the efficiency of the methods can be different [7-12].

## 3. Parallelization of the Statistical Algorithms

The parallelization of computations for the high-productive supercomputer systems appears to be one of the main ways of development of the modern computational mathematics. The supercomputers are the more and more widely used for a solution of the fundamental and applied problems in the areas of nuclear physics, climatology, economics, pharmacology, modeling of the training devices, and of the virtual reality, computational aerodynamics. Due to those specific features of the Monte Carlo methods, which were repeatedly stressed in the present paper, the statistical modeling begins to play the more and more noticeable role in all, indicated above areas of science and techniques. For these reasons, the actuality of the problems mentioned is growing very considerably, taking into account the fact that the computational aerodynamics is the most promoted area of the elaboration, development, and application of the Monte Carlo methods. As the mentioned above features of these methods permit to state, that the numerical schemes of a statistical modeling might be, in quite a natural way, transferred onto the parallel processors. The present authors are not aware of any studies on the parallelization of computations involving the methods, based on the modeling of trajectories of the "trial particles", but, nevertheless, the way of distribution of computations is, in this case, quite evident. Clearly, the successive modeling of the independent trajectories should be entrusted to the individual processors, while the information for the averaging will be gathered by a server.

Equally clearly is, that in this case, the productivity of the method is growing in direct proportionality to the number of parallel processors. As it is seen from the analysis of materials of the International Symposia on Rarefied Gas Dynamics, in the present time the statistical method, most popular all around the world, is the method of modeling of the "evolution of the ensemble of particles", proposed by Bird [13, 14], and the modifications of that method. Therefore, the most of scientific works on the distribution of computations is devoted just to that method.

One of the ways of such a distribution is, moreover, sufficiently evident. Since the structure of a computational algorithm possess a similarity for any temporal cut, at each of the processors the problem is set in its complete volume, while the information for averaging in time is coming to a server. Once again, it is evident, that the productivity of a method grows in a direct proportionality to the number of parallel processors.

Clearly, such an evident way of parallelization of the computations lays rather essential limitations on the class of those problems, which might be treated with the help of parallels computers. First, each individual processor should possess a sufficient power for the autonomous accommodation inside of itself of the complete problem, and such a demand is automatically excluding the modeling of complicated and multi-dimensional processes with a large volume of information. Second, lost is the possibility of using the memory of the parallel system in its totality, as well as its additional possibilities. And there is yet another reason, which stipulates the necessity for the elaboration of new methods of the distribution of computations. The supercomputers are rather expensive and not easily accessible. In the quality of their alternative appear the network cluster systems, or the network clusters, that is, the totalities of computers united by a network and possessing the common control. The network cluster systems are comparatively inexpensive, relatively simple in their arrangement, and by the sufficient quantity of active machinery they provide the productivity, comparable to that of supercomputers. Used here is the library of Message Passing Interface (MPI), which presents a low-level, but, at the same time, extremely convenient interface of the programming for a network cluster, and is based on the idea of exchange by communications among the parallel processors.

The programming for a network clusters is different from the usual models of programming on the basis of one processor, or even of the multitude of them. As concerns the network, the realization of the usual mechanism of the exchange of information is proved to be difficult because of the high overhead expenses, stipulated by the necessity to let to each of the processors the individual copy of one and the same dividable memory. Moreover, the solution of modern problems demands the increase of a total volume of the accessible operative memory of parallel computers. For this reason, by the programming for the network clusters is used, as a rule, the Single Program-Multiple Data technology (SPMD). The idea of SPMD consists in a tendency to divide the large array of information between the identical processors.

In this paper, we would like to describe the usage of *Message Passing Interface (MPI)*, which presents a low-level, but, at the same time, extremely convenient interface of the programming for a network cluster, and is based on the idea of exchange by communications among the parallel processors [8]. The programming for a network clusters is different from the usual models of programming on the basis of one

processor, or even of the multitude of them.

As concerns the network, the realization of the usual mechanism of the exchange of information is proved to be difficult because of the high overhead expenses, stipulated by the necessity to let to each of the processors the individual copy of one and the same dividable memory. Moreover, the solution of modern problems demands the increase of a total volume of the accessible operative memory of parallel computers. For this reason, by the programming for the network clusters is used, as a rule, the Single Program-Multiple Data technology (SPMD). The idea of SPMD consists in a tendency to divide the large array of information between the identical processors. After that each of the processors will carry out the processing of its part of data. In the case of SPMD-approach it would be sufficient to send from time to time to the processors the blocks of data, which are demanding the labor-consuming treatment, and afterwards — to gather the results of their work. If the time of processing of the block of data by one processor is considerably larger than the time of sending of that block through the network, then the network cluster systems becomes to be very effective. Just such an approach is used in MPI.

Here always exists a certain main processor, which realizes the distribution of data between the other processors, and upon the termination of computations gathers the results and demonstrates them to the user. Usually, after the distribution of data the master processor carries out the processing of the part of these data, thus aspiring to use the system's resources in the most effective way. Actually, each of the communications presents in itself a packet of the typified data, which one of the processors might send to the other one or to the group of them.

## 4. Structure of the Parallel Algorithm

Here in the basis of a parallel algorithm is set the method by Belotserkovskii–Yanitskii. The algorithm is elaborated taking into account the demand for the low expenditures on the acceptation/transmission and on the processing of information connected with the interaction of the master processor and the several subordinated processors. Thus, the structure of algorithm appears as follow:

1. The motion of particles is divided into two stages: collision and transfer.

2. Before the beginning of the calculations each of the subordinated processors obtains from the master one the range of those numbers of cells (the numeration of the cells is one-dimensional and through), which this processor should use, as well as the initial data introduced by user. After that each of the subordinated processors generates the particles within its own range of cells and in accordance to the initial conditions.

3. Begins the process of computation, by which each of the subordinated processors works only with its range of cells.

4. Before each of the new steps is carried out the computation of $dt$-step in time. Each of the subordinated processors sends its minimal $dt$ to the master one, where is determined the common minimal $dt$ and sent to all the subordinated processors.

5. Each of the subordinated processors realizes the collision of particles.

6. Each of the subordinated processors realizes the transfer of particles. During that event those particles, which fly out of the range of cell's numbers for the present processor, are transported into the buffer for the subsequent sending to the master processor.

7. Sending of the particles, which flew out of the range of cell's numbers to the master processor.

8. Sorting of the particles accepted into groups connected with processors and sending these particles to the subordinated processors.

9. Repeating of the points from 4 up to 8, till the moment, when will be reached the prescribed time $t$ of the experiment.

The scheme of the parallel algorithm working with a designation of the stages of the data exchange between the processors is as below:

*Master processor*

1. Introduction of the parameters of computation
2. Sending of the parameters of computation to the subordinated processor
3. Transference of the boundaries of a computational area
4. Minimal time step $dt$ to the master one
5. Obtainment and the choice of minimal step $dt$
6. Transference of the minimal temporal step $dt$
7. Sending the particles to the master processor
8. Obtainment of the arrays of particles
9. Formation of arrays for subordinate processors
10. Transfer of the arrays to the corresponding subordinate processors

Repetition of the steps from 5 to 10

*Subordinated processor*

1. Introduction of the parameters of computation
2. Obtainment of the parameters of computation
3. Obtainment of the boundaries of a computational area
4. Initialization of the initial distribution
5. Transference of the step $dt$
6. Obtainment and recording of step $dt$
7. The cycle over all the cells:

• Collision of the particles
• Transfer of the particles
• The particle in computational area array of the particles
• The particle out of computational area buffer for the transference

8. Transference of buffer to the master processor
9. Obtainment of the array of particles caught by the area of computations for the present computer
10. Repetition of the steps from 5 to 9

The use of a notion of the distributed operative memory, when the operative memory of several computers is acting as integrity. Such an approach is increasing not only the accessible volume of a memory, but also its aggregate working ability, since the processes of reading/writing are going on at several computers in parallel and independently of each other.

## 5. Some Results by Monte Carlo Methods

With the help of this method was solved, in particular, the rather labor-consuming problem of the interferention of two turbulent spots and aerodynamic characteristic of hypersonic vehicles [7, 8, 14, 15]. The turbulence presents in itself the most complicated area of the mechanics of fluid and gas, connected with the stochastical processes, with the complicated nonlinear equations, with the multi-dimensionality, and with the large volumes of information. The numerical study of the various phenomena in turbulence is hampered by the fact that the existing mathematical models of turbulence are not numerous and not perfect. Moreover, many of these phenomena do not possess the reliable physical model. For these reasons, those properties of the Monte Carlo methods in their application to the study of turbulent flows, which were formulated in the work [8, 16]. The distribution functions of turbulent spots are described in figure 1 and 2.

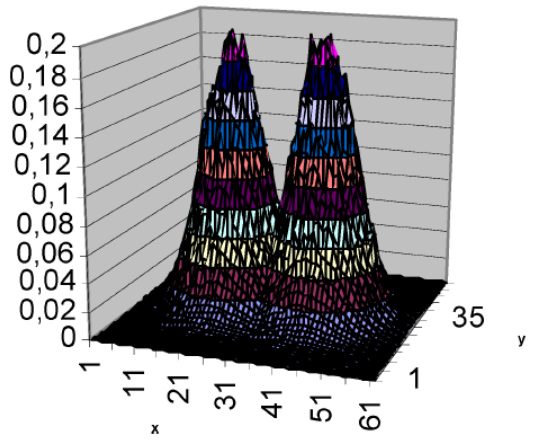


***Fig. 1.*** *The distribution functions of turbulent energy of the interacting spots at t = 0*

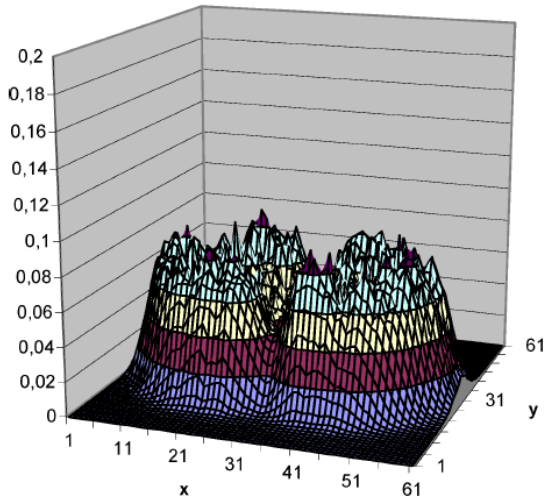


***Fig. 2.*** *Functions of the distribution of turbulent energy of the interacting spots at t = 40*

The DSMC method is commonly used to simulate rarefied flow problems, and the accuracy of the method depends directly on the accuracy of the gas-surface interaction model [9-12]. The Maxwell model is the most widely used and is based on classical thermodynamics in which it is assumed that molecules will either reflect diffusely from a surface with complete energy accommodation or will reflect specularly with no change in energy. An accommodation coefficient $\alpha_\tau$ is defined which specifies the fraction of molecules that will be scattered diffusely, with $\alpha_\tau = 0$ giving complete specular reflection, and $\alpha_\tau = 1$ giving complete diffuse reflection. In the (Cercignani-Lampic-Lord, CLL) model, the transformations of the normal and tangential components of velocity are assumed to be manually independent.

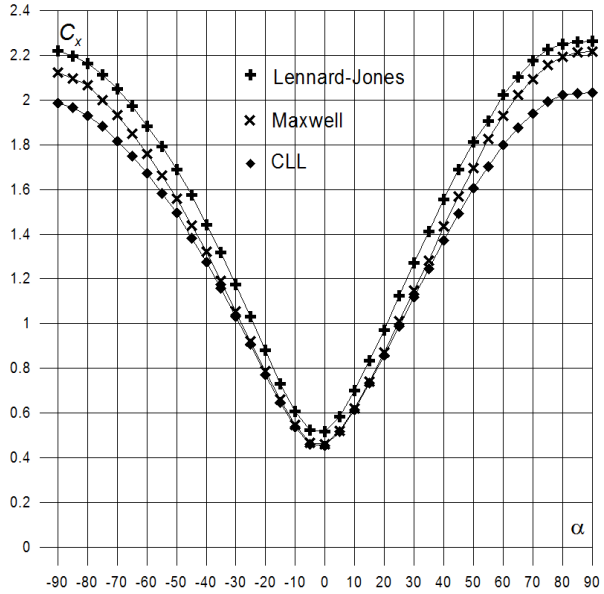


***Fig. 3.*** *Dependencies $C_x$ (α) with different gas-surface interaction models*

The Lennard-Jones potential function reflected the fact that the attractive forces dominated at high distances, and the repulsive forces dominated at small distances. The calculation was carried out in a range of angles of attack from –90° to 90°. The angle of attack changed by rotating the body around the center of mass around the axis z. Parameters of the problem as follows: velocity relationship $V_\infty / \sqrt{2RT_\infty} = 20$, the ratio of specific heat is 5/3. The total particles were using $5 \times 10^6$ particles. The tangential accommodation coefficient $\alpha_\tau = 1$ and the energy accommodation coefficient $\alpha_n = 1$.

The results of the aerodynamic characteristics of reentry vehicle by using Monte Carlo method with different gas-surface interaction models are described in fig 3. In figure shows the coefficient of aerodynamic force $C_x$ with application of different models (Maxwell, Lennard-Jones, CLL). The coefficient $C_x$ increase as the angle of attack increase. The multiple reflections have not been taken into account, since they are not significant for this body at the variation of the angle of attack. It is clear that the coefficient is sensitive to different models of interaction of molecules with surface. The method used in this paper is good for calculation aerodynamic characteristics of reentry vehicles.

## 6. Conclusions

With the connection Monte Carlo methods parallel computing techniques is growing the actuality of the studies conducted in many areas [17, 18], such as the computational aerodynamics of the most advanced area of physics and mathematics. Thus, formulated is the connection between the direct statistical modeling of the aerodynamical processes and the solution of kinetic equations, and it is shown that the contemporary stage of the development of computational methods proves to be inconceivable without a complex approach to the construction of algorithms taking into account all the peculiarities of the problem to be solved: the physical nature of a process, the mathematical model, the theoretical aspects of computational mathematics, and of stochastical processes. Considered are the possible ways of development of the methods of statistical modeling.

## Acknowledgments

The reported study was supported by Russian Science Foundation (research project No. 14-11-00709). Authors would like to thank Russian Science Foundation for their supports.

## References

[1] A. Hall, "On An Experimental Determination of π," Messenger Math, No. 2, 1873, pp. 113–114.

[2] N. Metropolis and S. Ulam, "The Monte-Carlo Method," J. Amer. Stat. Assoc, Vol. 44, No. 247, 1949, pp. 335–341.

[3] V.S. Vladimirov and I. M. Sobol, "Calculation of the Minimal Eigenvalue of the Peierls Equation by the Monte Carlo Method," Computational Mathematics and Mathematical Physics, 1958, No. 3, pp. 130–137.

[4] J.K. Haviland and M. D. Lavin, "Application of the Monte-Carlo Method to Heat Transfer in Rarefied Gases," Phys. Fluids, Vol. 5, No. 1, 1962, pp. 279–286.

[5] G.A. Bird, "Shock-Wave Structure in Rigid Sphere Gas," Rarefied Gas Dynamics, Academic, New York, Vol. 1, 1965, pp. 25–31.

[6] S.M. Ermakov, "Monte Carlo Method and Related Topics," Science, Moscow, 1985. (in Russian)

[7] Yu.I. Khlopkov,V.A. Zharov, and S.L. Gorelov, "Lectures on the Theoretical Methods of Study of Turbulence," Moscow, MPTI, 2005.

[8] O.M. Belotserkovskii and Y.I. Khlopkov, "Monte Carlo Methods in Mechanics of Fluid and Gas," World Scientific Publishing Co. New Jersy, London, Singapore, Beijing, Hong Kong, 2010.

[9] O.M. Belotserkovskii, "Numerical Simulation in Mechanics of Continua," Moscow, 1984.

[10] Zay Yar Myo Myint and A.Yu. Khlopkov, "Aerodynamic Characteristics of an Aircraft with a Complex Shape Taking into Account the Potential of Molecular Flow Interaction with a Surface," TsAGI Science Journal, Vol. 41, No. 5, 2010, pp. 551-566.

[11] Zay Yar Myo Myint, Yu.I. Khlopkov and A.Yu. Khlopkov, "Aerothermodynamics Investigation for Future Hypersonic Aerospace Systems," Proceedings of the 4th International Conference on Science and Engineering, Yangon, Myanmar, 9-10 December, 2013. (CD ROM)

[12] Yu.I. Khlopkov, S.L. Chernyshev, Zay Yar Myo Myint and A.Yu. Khlopkov, "Introduction to Speciality II. High-Speed Aircrafts," Moscow, MIPT, 2013.

[13] G. A. Bird, "Molecular Gas Dynamics and the Direct Simulation of Gas Flows," Oxford University Press, 1994.

[14] Yu.I. Khlopkov, "Statistical modeling in computational aerodynamics," Moscow, MIPT, 2006.

[15] Yu.I. Khlopkov, Tian Van Vyong, Zay Yar Myo Myint and A.Yu. Khlopkov, "Monte-Carlo methods in applied mathematics and computational aerodynamics," Proceeding of 29th congress of the international council of the aeronautical sciences, St. Petersburg, September 7-12, 2014.

[16] A.Yu.Khlopkov, Zay Yar Myo Myint, Yu.I. Khlopkov, "Parallel computing system of Monte Carlo Methods," Materials of the international conference Fundamental researches, Dominican Republic, April 13-22, 2014, pp. 40-41.

[17] A.Yu.Khlopkov, Zharov V.A., Khlopkov A.Yu., and Zay Yar Myo Myint, "Development of Monte Carlo Methods in Hypersonic Aerodynamics," Universal Journal of Physics and Application, Vol. 2, No. 4, 2014, pp. 213-220.

[18] Fan Bai, "Parallel Monte Carlo Methods," Georgia State University.
http://www.cs.gsu.edu/~fbai1/csc8530/Survey_Bai.pdf