Statistical errors of DSMC results for rarefied gas flows

S.V. Rogasinsky*, D.A. Levin† and M.S. Ivanov**

*Institute of Computational Mathematics and Mathematical Geophysics SB RAS, 630090, Novosibirsk, Russia
†Pennsylvania State University, University Park, PA 16802, USA
**Institute of Theoretical and Applied Mechanics SB RAS, 630090, Novosibirsk, Russia

Abstract. In the article the statistical errors of the numerical results received by DSMC method are considered. These errors are connected with statistical nature of DSMC schemes. Statistical errors in the DSMC method were analyzed previously for a particular case of rarefied gas dynamics problems with zero mean velocity. Based on the general theory of the Monte Carlo methods, a detailed analysis of the variance of statistical estimates used in the DSMC method for calculating physical characteristics of rarefied gas flows in the general case was performed in the present work. The practical manner for application of a variance calculation is offered.

Keywords: DSMC method, statistical error, variance
PACS: 02.70.Uu, 02.60.-x, 02.70.-c

INTRODUCTION

The Direct Statistical Monte Carlo (DSMC) method is the main numerical method for solving multidimensional problems of rarefied gas dynamics [1, 2]. Flexibility of the method allows natural allowance for various physical and chemical processes that occur in binary collisions of particles, such as, e.g., excitation of internal degrees of freedom of molecules and chemical reactions. In addition, because of comparative simplicity of modeling various boundary conditions and complicated geometry in the DSMC method, the latter is often the only acceptable numerical method for solving problems of applied high-altitude aerothermodynamics.

Nevertheless, the issue of statistical errors of DSMC simulation of such problems, which is of great importance for practice, has not been adequately addressed. Statistical errors in the DSMC method were analyzed previously in [3] , where a particular case of rarefied gas dynamics problems with zero mean velocity was considered. Based on the general theory of the Monte Carlo methods [4], a detailed analysis of the variance of statistical estimates used in the DSMC method for calculating physical characteristics of rarefied gas flows in the general case was performed in the present work. For simplicity, a simple monatomic gas was considered.

ASYMPTOTIC EXPRESSION OF VARIANCE FOR NUMBER DENSITY ESTIMATOR

In stationary case physical characteristics of rarefied gas dynamic, which calculated by the DSMC method, can be presented by mathematical expectation $E[Z(X)]$ of any random variable $Z(X)$, or by ratio of two mathematical expectations of two random variable $E[Z(X)]/E[Y(X)]$. Here X is state of N particle system at time moment t, which charges by discrete step Δt. The mathematical expectations can be estimated by simulating states of N particle system X_1, X_2, X_M consistently i.e.

$$Z_M = \frac{1}{M}[Z(X_1) + Z(X_2) + + Z(X_M)].$$

If $E[Z^2(X)] < \infty$ random error of this formula can be estimated by center limit theorem

$$P\{|E[Z(X)] - Z_M| < \alpha \cdot \frac{\sigma(Z)}{\sqrt{M}}\} = erf(\alpha), \qquad \sigma^2(Z) = D\{Z\} = E[Z^2(X)] - (E[Z(X)])^2,$$

where $E[Z^2(X)]$ can be estimated in simulating process of the N particle system states.

Consider following physical values.

$$n = \int f(\mathbf{r},\mathbf{v})d\mathbf{v}; \quad nu_\alpha = \int v_\alpha f(\mathbf{r},\mathbf{v})d\mathbf{v}; \quad 3nkT = m\int \mathbf{v}^2 f(\mathbf{r},\mathbf{v})d\mathbf{v} - mn\mathbf{u}^2;$$

$$P_{\alpha\beta} = m\int v_\alpha v_\beta f(\mathbf{r},\mathbf{v})d\mathbf{v} - mnu_\alpha u_\beta;$$

$$q_\alpha = \frac{m}{2}\int \mathbf{v}^2 v_\alpha f(\mathbf{r},\mathbf{v})d\mathbf{v} - \frac{m}{2}u_\alpha \int \mathbf{v}^2 f(\mathbf{r},\mathbf{v})d\mathbf{v} - \sum_{\beta=1}^{3} P_{\alpha\beta}u_\beta;$$

Here and in the further indexes α, β run from 1 to 3 and $\mathbf{v} = (v_1, v_2, v_3)$. Easy note that these values are functions from moments of function $f(\mathbf{r},\mathbf{v})$, and hence they can be presented by mathematical expectations of random variables :

$$\xi_0 = \sum_{i=1}^{N} h(\mathbf{r}_i); \quad \xi_1^\alpha = \sum_{i=1}^{N} h(\mathbf{r}_i)(\mathbf{v}_i)_\alpha; \quad \xi_2^{\alpha\beta} = \sum_{i=1}^{N} h(\mathbf{r}_i)(\mathbf{v}_i)_\alpha(\mathbf{v}_i)_\beta; \quad \xi_3^\alpha = \sum_{i=1}^{N} h(\mathbf{r}_i)\mathbf{v}_i^2(\mathbf{v}_i)_\alpha,$$

for which

$$E(\xi_0) = \int\int h(\mathbf{r})f(\mathbf{r},\mathbf{v})d\mathbf{r}d\mathbf{v}, \quad E(\xi_1^\alpha) = \int\int h(\mathbf{r})v_\alpha f(\mathbf{r},\mathbf{v})d\mathbf{r}d\mathbf{v},$$

$$E(\xi_2^{\alpha\beta}) = \int\int h(\mathbf{r})v_\alpha v_\beta f(\mathbf{r},\mathbf{v})d\mathbf{r}d\mathbf{v}, \quad E(\xi_3^\alpha) = \int\int h(\mathbf{r})\mathbf{v}^2 v_\alpha f(\mathbf{r},\mathbf{v})d\mathbf{r}d\mathbf{v}.$$

where $h(\mathbf{r})$ is characteristic function of cell. Starting from that physical values $n(\mathbf{r}), \quad u_\alpha(\mathbf{r}), \quad T(\mathbf{r}), \quad P_{\alpha\beta}(\mathbf{r}), \quad q_\alpha(\mathbf{r})$ are presented in forms

$$\tilde{n}(\mathbf{r}) = \frac{1}{\Delta\mathbf{r}} E(\xi_0); \quad \tilde{u}_\alpha(\mathbf{r}) = \frac{E(\xi_1^\alpha)}{E(\xi_0)}; \quad \frac{3k}{m}\tilde{T}(\mathbf{r}) = \sum_{\alpha=1}^{3} \frac{E(\xi_2^{\alpha\alpha})}{E(\xi_0)} - \sum_{\alpha=1}^{3}\left(\frac{E(\xi_1^\alpha)}{E(\xi_0)}\right)^2;$$

$$\frac{1}{m}\tilde{P}_{\alpha\beta}(\mathbf{r}) = \frac{1}{\Delta\mathbf{r}} E(\xi_2^{\alpha\beta}) - \frac{1}{\Delta\mathbf{r}} E(\xi_1^\alpha)\cdot\frac{E(\xi_1^\beta)}{E(\xi_0)};$$

$$\frac{2}{m}\tilde{q}_\alpha(\mathbf{r}) = \frac{1}{\Delta\mathbf{r}} E(\xi_3^\alpha) - \frac{E(\xi_1^\alpha)}{E(\xi_0)}\cdot \frac{1}{\Delta\mathbf{r}} \sum_{\beta=1}^{3} E(\xi_2^{\beta\beta}) - 2\cdot\sum_{\beta=1}^{3}(\frac{1}{m}\tilde{P}_{\alpha\beta})\cdot\frac{E(\xi_1^\beta)}{E(\xi_0)}.$$

Statistical estimators of values which are used in the DSMC method are obtained by exchange of corresponding mathematical expectations to their statistical estimators, i.e.

$$E[\zeta] \approx \hat{\zeta} = \frac{1}{M}\sum_{s=1}^{M}\zeta_s, \tag{1}$$

where M is sample size, ζ_i is magnitude of random variable at N particle system state X_i. Thus statistical estimator of values (1) have forms $\hat{Q} = \tilde{Q}(\hat{\zeta}_1,\ldots,\hat{\zeta}_p)$. and $\tilde{Q}(x_1,\ldots,x_p)$ is smooth function.

For example, the following statistical estimator is used for calculation of mean velocity components

$$\hat{u}_\alpha = \frac{\frac{1}{M}\sum_{s=1}^{M}(\xi_1^\alpha)_s}{\frac{1}{M}\sum_{s=1}^{M}(\xi_0)_s.} \tag{2}$$

Here $\tilde{Q}(x_1,x_2) = \frac{x_1}{x_2}$. Take $x_1 = \frac{1}{M}\sum_{s=1}^{M}(\zeta_1^{(\alpha)})_s$, $x_2 = \frac{1}{M}\sum_{s=1}^{M}(\zeta_0)_s$, we receive estimator (2).

In the further analysis of variations which are obtained by this manner is carried out.

It is well-known from general theory of Monte Carlo method the statistical estimator $\tilde{Q}(\hat{\zeta}_1,\ldots,\hat{\zeta}_p)$ is asymptotic unbiased estimator:

$$E(\hat{Q}) = E\{\tilde{Q}(\hat{\zeta}_1,\ldots,\hat{\zeta}_p)\} \to \tilde{Q}(E(\zeta_1),\ldots,E(\zeta_p)) \qquad \text{at} \quad M\to\infty$$

and its variance is asymptotic form

$$D\{\hat{Q}\} = \frac{1}{M}\sum_{i=1}^{p}\sum_{j=1}^{p} cov\{\zeta_i\zeta_j\}F_iF_j, \tag{3}$$

where $cov\{\zeta_i\zeta_j\} = E\{\zeta_i\zeta_j\} - E\{\zeta_i\}\cdot E\{\zeta_j\}$, $i,j = 1,\ldots,p$ and $F_i = \frac{\partial}{\partial x_i}Q(x_1,\ldots,x_p)$ is derivative at values $x_i = E(\zeta_i)$, $i = 1,\ldots,p$. Since random variables ζ_i and ζ_j have forms $\zeta_i = \sum_{k=1}^{N} h(\mathbf{r}_k)\varphi_i(\mathbf{v}_k)$ and $\zeta_j = \sum_{k=1}^{N} h(\mathbf{r}_k)\varphi_j(\mathbf{v}_k)$, then

$$cov\{\zeta_i\zeta_j\} = N\cdot[E\{h(\mathbf{r}_1)\varphi_i(\mathbf{v}_1)\varphi_j(\mathbf{v}_1)\} - E\{h(\mathbf{r}_1)\varphi_i(\mathbf{v}_1)\}\cdot E\{h(\mathbf{r}_1)\varphi_j(\mathbf{v}_1)\}]+$$

$$+N(N-1)\left[E\{h(\mathbf{r}_1)h(\mathbf{r}_2)\varphi_i(\mathbf{v}_1)\varphi_j(\mathbf{v}_2)\} - E\{h(\mathbf{r}_1)\varphi_i(\mathbf{v}_1)\}\cdot E\{h(\mathbf{r}_1)\varphi_j(\mathbf{v}_1)\}\right],$$

here

$$E\{h(\mathbf{r}_1)\varphi_i(\mathbf{v}_1)\} = \int h(\mathbf{r}_1)\varphi_i(\mathbf{v}_1)f_N^{(1)}(\mathbf{r}_1,\mathbf{v}_1)d\mathbf{r}_1 d\mathbf{v}_1,$$
$$E\{h(\mathbf{r}_1)h(\mathbf{r}_2)\varphi_i(\mathbf{v}_1)\varphi_j(\mathbf{v}_2)\} = \int h(\mathbf{r}_1)\varphi_i(\mathbf{v}_1)h(\mathbf{r}_2)\varphi_j(\mathbf{v}_2)f_N^{(2)}(\mathbf{r}_1,\mathbf{v}_1,\mathbf{r}_2,\mathbf{v}_2)d\mathbf{r}_1 d\mathbf{v}_1 d\mathbf{r}_2 d\mathbf{v}_2.$$

These general formulas which are received above solve question of calculation variances for estimator $\hat{Q}(...)$.

We will use condition, which is imposed on cell size $\Delta\mathbf{r}$. This condition is $\Delta\mathbf{r}\to 0$. Passing to limit in formula for variance we obtain $\lim\limits_{\Delta\mathbf{r}\to 0}\Delta\mathbf{r}\cdot D\{\hat{Q}\} = \frac{1}{M}\sum\limits_{i=1}^{p}\sum\limits_{j=1}^{p} c_{ij}G_iG_j$, where $c_{ij} = \lim\limits_{\Delta\mathbf{r}\to 0}\frac{cov\{\zeta_i\zeta_j\}}{\Delta\mathbf{r}}$ and $G_i = \lim\limits_{\Delta\mathbf{r}\to 0}\Delta\mathbf{r}\cdot F_i$. We easy calculate magnitude c_{ij} at $\Delta\mathbf{r}\to 0$.

$$c_{ij} = \lim_{\Delta\mathbf{r}\to 0} N\cdot\frac{E\{h(\mathbf{r}_1)\varphi_i(\mathbf{v}_1)\varphi_j(\mathbf{v}_1)\}}{\Delta\mathbf{r}} = \int\varphi_i(\mathbf{v})\varphi_j(\mathbf{v})f(\mathbf{r},\mathbf{v})d\mathbf{v}.$$

Finally we obtain $D\{\hat{Q}\}\sim\frac{1}{\Delta\mathbf{r}}\frac{1}{M}\sum\limits_{i=1}^{p}\sum\limits_{j=1}^{p} c_{ij}G_iG_j$. Using inequality Cauchy-Bunyakovsky we have

$$|\sum_{i=1}^{p}\sum_{j=1}^{p} c_{ij}G_iG_j| \le \sum_{i=1}^{p}\int\varphi_i^2(\mathbf{v})f(\mathbf{r},\mathbf{v})d\mathbf{v}\cdot\sum_{j=1}^{p}G_j^2,$$

then

$$D\{\hat{Q}\}\sim\frac{1}{\Delta\mathbf{r}}\frac{1}{M}\sum_{i=1}^{p}\int\varphi_i^2(\mathbf{v})f(\mathbf{r},\mathbf{v})d\mathbf{v}\cdot\sum_{j=1}^{p}G_j^2 \qquad (3').$$

This expression for variance of random variable $\hat{Q}(...)$ is upper bound, but it has preference compared with previous expression because for calculation it only requires diagonal elements of matrix c_{ij}.

Asymptotic expression of variance for number density estimator.

Statistical estimator for calculation of number density has form $\hat{n} = \frac{1}{\Delta\mathbf{r}}\cdot\frac{1}{M}\cdot\sum\limits_{s=1}^{M}(\xi_0)_s$, where $\xi_0 = \sum\limits_{i=1}^{N}h(\mathbf{r}_i)$. Variance of this statistical estimator is

$$D\{\hat{n}\}\sim\frac{1}{\Delta\mathbf{r}}\cdot\frac{1}{M}\cdot n(\mathbf{r}).$$

Asymptotic expression of variance for mean velocity components.

Statistical estimator for calculation of number density has form (2). Variance of this statistical estimator is

$$D\{\hat{u}_\alpha\}\sim\frac{1}{\Delta\mathbf{r}}\cdot\frac{1}{M}\cdot\frac{\langle v_\alpha^2\rangle - u_\alpha^2}{n(\mathbf{r})}.$$

Asymptotic expression of variance for temperature.

Statistical estimator for calculation of temperature is received from expression

$$3\tilde{T}(\mathbf{r}) = \sum_{\alpha=1}^{3}\frac{E(\xi_2^{\alpha\alpha})}{E(\xi_0)} - \sum_{\alpha=1}^{3}\left(\frac{E(\xi_1^{\alpha})}{E(\xi_0)}\right)^2.$$

Factors of variance of this statistical estimator are $\mathbf{G} = \frac{1}{n(\mathbf{r})}(\mathbf{u}^2 - 3T\ ,1\ ,1\ ,1\ ,-2u_1\ ,-2u_2\ ,-2u_3)$, and sum of diagonal elements of matrix c_{ij} is

$$\sum_{i=1}^{7}c_{ii} = n(\mathbf{r})\cdot(1+\langle v_1^2\rangle+\langle v_2^2\rangle+\langle v_3^2\rangle+\langle v_1^4\rangle+\langle v_2^4\rangle+\langle v_3^4\rangle).$$

For local-maxwellian distribution on velocities we have

$$\sum_{i=1}^{7}c_{ii} = n(\mathbf{r})(\frac{23}{2}+4\mathbf{U}^2+U_1^4+U_2^4+U_3^4).$$

Asymptotic expression of variance for stress tensor components.

Statistical estimator for calculation of stress tensor components are received from expression

$$\tilde{P}_{\alpha\beta}(\mathbf{r}) = \tfrac{1}{\Delta\mathbf{r}}\, E(\xi_2^{\alpha\beta}) - \tfrac{1}{\Delta\mathbf{r}}\, E(\xi_1^{\alpha}) \cdot \frac{E(\xi_1^{\beta})}{E(\xi_0)}.$$

Factors of variance of this statistical estimator are $\mathbf{G} = (u_\alpha \cdot u_\beta\ , -u_\beta\ , -u_\alpha\ , 1)$, and sum of diagonal elements of matrix c_{ij} is $\sum\limits_{i=1}^{4} c_{ii} = n(\mathbf{r}) \cdot (1 + \langle v_\alpha^2 \rangle + \langle v_\beta^2 \rangle + \langle v_\alpha^2 v_\beta^2 \rangle)$.

For local-maxwellian distribution on velocities we have

$$\sum_{i=1}^{4} c_{ii} = n(\mathbf{r})(3 + 2\delta_{\alpha\beta} U_\alpha^2 + \frac{3}{2}(U_\alpha^2 + U_\beta^2) + U_\alpha^2 U_\beta^2).$$

Asymptotic expression of variance for energy vector components.

Statistical estimator for calculation of stress tensor components are received from expression

$$2\tilde{q}_\alpha(\mathbf{r}) = \tfrac{1}{\Delta\mathbf{r}}\, E(\xi_3^{\alpha}) - \frac{E(\xi_1^{\alpha})}{E(\xi_0)} \cdot \tfrac{1}{\Delta\mathbf{r}} \sum_{\beta=1}^{3} E(\xi_2^{\beta\beta}) - 2 \cdot \sum_{\beta=1}^{3} (\frac{1}{m}\tilde{P}_{\alpha\beta}) \cdot \frac{E(\xi_1^{\beta})}{E(\xi_0)}.$$

Factors of variance of this statistical estimator are

$$\mathbf{G} = (3\langle \mathbf{v}^2 \rangle u_\alpha + 2 \sum_{\beta \neq \alpha}^{3} \langle v_\alpha v_\beta \rangle u_\beta - 4\mathbf{u}^2 u_\alpha\ , 4u_\alpha^2 + 2\mathbf{u}^2 - 3\langle \mathbf{v}^2 \rangle, 4u_\alpha u_{\beta_1} - 2\langle v_\alpha v_{\beta_1} \rangle, 4u_\alpha u_{\beta_2} - 2\langle v_\alpha v_{\beta_2} \rangle,$$

$$-3u_\alpha, -3u_\alpha, -3u_\alpha, -2u_{\beta_1}, -2u_{\beta_2}, 1).$$

and sum of diagonal elements of matrix c_{ij} is

$$\sum_{i=1}^{10} c_{ii} = n(\mathbf{r}) \cdot (1 + \langle \mathbf{v}^2 \rangle + \langle v_1^4 \rangle + \langle v_2^4 \rangle + \langle v_3^4 \rangle + \langle \mathbf{v}^4 v_\alpha^4 \rangle + \langle v_\alpha^2 \mathbf{v}^2 \rangle - \langle v_\alpha^4 \rangle).$$

For local-maxwellian distribution on velocities we have

$$\sum_{i=1}^{10} c_{ii} = n(\mathbf{r})(\frac{97}{2} + \frac{37}{2}\mathbf{U}^2 + U_1^4 + U_2^4 + U_3^4 + \frac{127}{2}U_\alpha^2 + 10U_\alpha^2\mathbf{U}^2 + \frac{1}{2}\mathbf{U}^4 - U_\alpha^4 + \mathbf{U}^4 U_\alpha^2).$$

From explicit expression of variance for random variable $\hat{Q}$ it is followed the asymptotic expression for the dimensionless root-mean-square error for a random quantity Q, which is the estimate of some basic hydrodynamic quantity, turned out to have a simple form

$$\frac{\sigma(\hat{Q})}{|E(\hat{Q})|} = \frac{\sqrt{D\{\hat{Q}\}}}{|E(\hat{Q})|} \sim \frac{1}{\sqrt{M}} \frac{C(\mathbf{r})}{\sqrt{\Delta\mathbf{r} n(\mathbf{r})}},$$

where M is the sampling volume, $n(\mathbf{r})$ is the number density at the point $\mathbf{r}$, and $\Delta\mathbf{r}$ is the cell volume.

CALCULATION OF VARIANCE IN APPLICATIONS OF THE DSMC METHOD

Manner for calculation of variance by batch.

Mathematical expectation $E[Z(X)]$ random variable $Z(X)$ is estimated by formula

$$E[Z(X)] \approx \zeta_M = \frac{1}{M}\sum_{i=1}^{M} Z_i, \qquad Z_i = Z(X_i).$$

We regroup terms in right side of this expression selecting batchs of values of random variable Z_i. Let be $M = L \cdot m$, here m is size of batch and L is number of batch.

$$\zeta_M = \frac{1}{M}\sum_{i=1}^{M} Z_i = \frac{1}{L}\sum_{i=1}^{L} (\frac{1}{m}\sum_{k=(i-1)m+1}^{(i-1)m+m} Z_k) = \frac{1}{L}\sum_{i=1}^{L} \hat{Z}_i, \qquad \hat{Z}_i = \frac{1}{m}\sum_{k=(i-1)m+1}^{(i-1)m+m} Z_k.$$

Obviously this estimator is unbiased, i.e. $E[\zeta_M] = E[Z(X)]$. Variance of random variable is

$$D[\zeta_M] = \frac{1}{L} D[\hat{Z}] \approx \frac{1}{L}\Big[\frac{1}{L}\sum_{i=1}^{L} \hat{Z}_i^2 - \Big(\frac{1}{L}\sum_{i=1}^{L} \hat{Z}_i\Big)^2\Big] \tag{4}$$

If we take value $m = 1$ from this expression is receives usual estimator of variance for random variable $Z(X)$.

In figure 1,2 results of comparison of the variance behaviors for density, temperature and speed for a classical problem of heat transfer between two parallel plates are present. Model of molecules was hard spheres and Knudsen number was 0.1. The ratio of the plate temperatures was equal to 4. Calculations were made by the DSMC method with use of formulas $(3')$ (a continuous curve) and (4) (a curve of points). The top curves in figures show profiles of the corresponding values variances on one sampling. The lower lines are corresponding mean square deviation σ. From figures it is visible, that in the given problem of the variance calculations under formulas $(3')$ and (4) yield about identical results, and the variance calculated by asymptotic formula $(3')$ as follows from Figure 2, the value of the variance is overstated in comparison with formula (4). For a practical estimation of variance of calculated values in the DSMC method it can recommend to use formula (4) because it has less capacity.

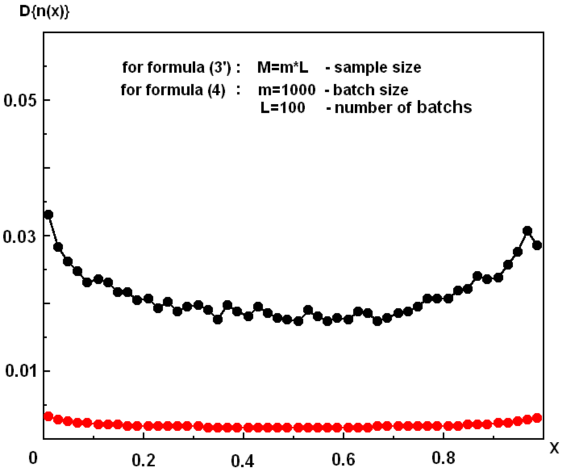

FIGURE 1.

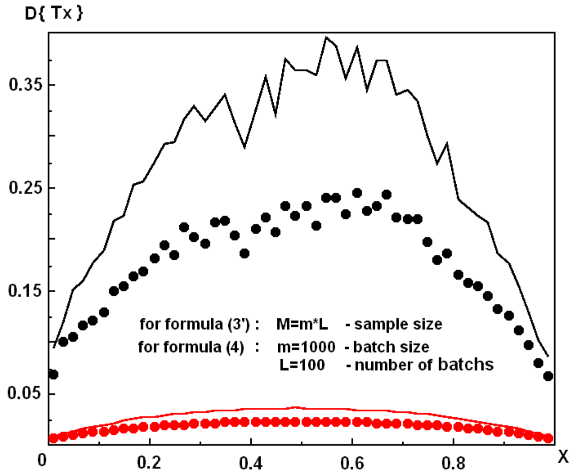

FIGURE 2.

REFERENCES

1. M.S. Ivanov, S.F. Gimelshein, Computational hypersonic rarefied flows. Ann. Rev. Fluid Mech. 1998, Vol. 30, pp. 469-505.
2. M.S. Ivanov, G.N. Markelov and S.F. Gimelshein, Statistical simulation of reactive rarefied flows: numerical approach and applications, AIAA Paper 98-2669, June 1998.
3. D.A. Fedosov, S.V. Rogasinsky, M.I. Zeifman, M.S. Ivanov, A.A. Alexeenko, and D.A. Levin, Analysis of numerical errors in the DSMC method. 24th International symposium on rarefied gas dynamics, Monopoli (Bari),Italy, 10-16 July 2004, pp. 589-594
4. S. M. Ermakov, G. A. Mikhailov, Course of Statistical Modeling, Nauka, Moscow, 1984.