A Probabilistic Automatic Steady State Detection Method for the Direct Simulation Monte Carlo

A. Karchani, O. Ejtehadi, and R. S. Myong

Department of Aerospace and System Engineering, Research Center for Aircraft Parts Technology, Gyeongsang National University, Jinju, Gyeongnam 660-701, South Korea.

myong@gnu.ac.kr

Abstract: The statistical error associated with sampling in the DSMC can be categorized as type I and II, which are caused by the incorrect rejection and acceptance of the null hypothesis, respectively. In this study, robust global and local automatic steady state detection methods were developed based on an ingenious method based purely on the statistics and kinetics of particles. The key concept is built upon probabilistic automatic reset sampling (PARS) to minimize the type II error caused by incorrect acceptance of the samples that do not belong to the steady state. The global steady state method is based on a relative standard variation of collisional invariants, while the local steady state method is based on local variations in the distribution function of particles at each cell. In order to verify the capability of the new methods, two benchmark cases—the one-dimensional shear-driven Couette flow and the two-dimensional high speed flow past a vertical wall—were extensively investigated. Owing to the combined effects of the automatic detection and local reset sampling, the local steady state detection method yielded a substantial gain of 30-36% in computational cost for the problem studied. Moreover, the local reset feature outperformed the automatic detection feature in overall computational savings.

AMS subject classifications: 82B80, 76P05

Keywords: Direct simulation Monte Carlo, statistical error, type II error, steady state detection, reset sampling.

1. Introduction

The direct simulation Monte Carlo (DSMC) method is considered as one of the most successful computational methods to solve the Boltzmann equation based on direct statistical simulation of the molecular processes described by the gas kinetic theory [1, 2]. It has also been mathematically and empirically proved that the DSMC solution will converge to the true solution of the Boltzmann equation for a gas undergoing binary collisions between gas particles, if critical computational parameters—time-step, cell-size, and the number of particles—are chosen properly and when no wall surface boundary condition is involved in the simulation [3, 4]. Owing to robustness and ease of incorporating various collision mechanisms into the algorithm, DSMC has expanded its way into diverse applications including hypersonic gas flows, micro-scale gases, chemical reactions, and material processing [5-8].

Four types of computational error are in general present in the DSMC simulation [4]: decomposition (or discretization), statistical [9, 10], machine (or round-off), and boundary condition errors, as summarized in Fig. 1. In particular, the statistical error—the focus of the present work—is caused by the random fluctuations and statistical uncertainty inherent in the DSMC method. In a recent study of the verification method for DSMC [4], based on the exact physical laws of conservation, it was shown that the statistical error is dominant in the first phase of error convergence and the rate of its decrease is inversely proportional to the square root of the sample steps. In the second phase, the combination of boundary condition and decomposition errors becomes prominent in comparison with the statistical error. To reduce the statistical uncertainty and noise arising in evaluating the mean value of the random variables during the DSMC simulation, an appropriate probability sampling process is required [1, 4, 11].

The statistical error associated with the sampling procedure can be further categorized as type I and type II. In order to statistically analyze these errors, a null hypothesis [12] is defined as the DSMC samples belonging to the steady state. An *incorrect rejection* of the null hypothesis—not considering the samples that belong to the steady state—leads to *type I error*, which can be minimized by including more independent steady state samples. On the other hand, an *incorrect acceptance* of

the null hypothesis—considering the samples that do not belong to the steady state—leads to *type II error*, which can be minimized by accurately detecting the onset of the steady state and by starting sampling after completely reaching the steady state.

Minimization of the type I error was already investigated in many previous studies on the DSMC [4, 13-18]. However, very few efforts have been directed to the type II error in the DSMC simulation. When the system properties vary continuously, sampling in the unsteady phase can lead to a wrong contribution to the steady state solution of the DSMC simulation. Therefore, it is of critical importance to determine precisely when to initiate a sampling for a steady state solution. Consequently, correctly identifying the termination of the unsteady phase, which is directly related to the type II error, can be used as the onset criterion for starting a sampling. This necessitates the concept of restarting sampling automatically in the DSMC simulation.

In traditional DSMC methods, the number of time steps required to reach a steady state is specified by user as an input parameter [1, 19]. This adjustable parameter is basically estimated by considering the trade-off between accuracy and efficiency. However, designating more time steps will demand higher computational cost while premature initiation of sampling will compromise the accuracy due to the type II error. Hence, the conventional methods relying completely on the user's input are not effective, since predicting the exact number of time steps to start a sampling is not only difficult but also highly problem-dependent. Moreover, the drawback of problem-dependency persists in an alternative method of estimating the number of required time steps by multiplying a constant value to the time required for the thermal information to traverse the simulated flow domain by means of acoustic waves [20].

To overcome this shortfall, Bird [21, 22] proposed an automatic method for estimating steady state convergence on the basis of the variation of total number of simulated particle (i.e., the first component of the Boltzmann collisional invariants) throughout the simulation domain. However, the effect of variation of higher order collisional invariants (momentum and kinetic energy) were not considered, which led to inability of the method in the case of constant number of particles or in presence of highly unsteady regions. Burt and Boyd [23] also proposed a method that tackles insensitivity of the Bird's method to weak transient behavior in small regions of relatively low density

or recirculating flow. The method was mainly based on the variation of particle fluxes on the boundaries at two successive time steps, and introduction of the convergence criterion integrating the particle flux differences variance. The method, though benefited some saving in computational cost and simplicity of implementation, was found to suffer from the problem dependency and inability to deal with problems allowing scatter-induced fluctuations in the location of high gradient regions such as oblique shocks or shock-boundary layer interaction [23].

At present, a number of important issues still remain unsolved regarding developing an accurate and efficient steady state detection method. The main drawback of the existing non-automatic methods is that a great deal of knowledge and experience on the DSMC method and flow behavior are required in order to determine the exact number of time steps needed to start a sampling. Moreover, the previous automatic methods faces important drawbacks: the impossibility of application to problems with a constant number of particles in the domain, such as one dimensional (1-D) Couette flow, Fourier heat flow, and shock wave structure. Further, the method was ineffective when the flow characteristics were still evolving after the total number of particles had already reached an asymptotic value. Such situation can be found in external flows around an object with recirculating regions, flows containing separation regions, or highly localized regions involving a large variation in the density of gas.

In this study, we present a probabilistic automatic steady state detection method on the basis of the statistical features of the gas particles that include the effect of the variation of the momentum and energy along with the number of particles. The method can be easily implemented in the current DSMC algorithms and applied to handle local recirculating flows or low energy regions without any difficulty.

Furthermore, since each cell can be considered as an independent statistical sample cell in the new method, an effective strategy to apply the sampling process—in terms of the number of samples and times the sample is reset—specifically to local cells can be developed. This will be called variable sampling procedure or *local* steady state detection method. Note that the number of particles within the cell multiplied by the number of samples taken for the cell—the sample size—is a unique property of each cell which itself is an independent statistical population. In the previous methods, the idea of

using a different value of sample size to each cell was not fully developed, which led to significant degradation in computational efficiency. The main features, advantages, and drawbacks of the existing and new steady state detection methods are summarized in table 1 for a detailed comparison. Finally, in order to confirm the capability of the new method, two cases—the 1-D shear-driven Couette flow and the two-dimensional (2-D) high speed flow past a vertical wall—were extensively investigated.

2. Steady State Detection Method

2.1. Phase portrait analysis of the DSMC sampling

In order to compute the macroscopic properties accurately, an efficient implementation of the probabilistic sampling process in the DSMC algorithm is essential. As explained in the Introduction, the type II error arising from incorrect acceptance of the samples that do not belong to steady state can be minimized by correctly identifying the termination of unsteady phase. Several possible statistical approaches for analyzing sampling process are available. Here, phase portraits of the sampling estimators and distribution of samples are considered to investigate the effect of incorrect inclusion of the samples that do not belong to steady state in sampling procedure.

The phase diagrams of variables used in the DSMC sampling procedure are illustrated for a system (1-D Couette flow) with 2,000 particles per cell in Fig. 2. For the case in which variables are correlated, the cloud tends to form an elliptical shape as seen in Fig. 2 (b). Moreover, the portraits in phase (C_x, C_y) , (C_x^2, C_x) , (C_x^2, C_y^2) show the evolution of system from initial state (marked in the figure as 'start point') to steady state. The points forming a cloud may be used as an indicator of steady state. For example, unsteady transition from initial condition is apparent in the phase portraits. In particular, the higher order phase portrait defined in the squared particle velocities (C_x^2, C_y^2) in Fig. 2 (c) provides the best illustration of transition phase.

The phase portraits for different number of particles per cell are plotted in Fig. 3. As the number of particles decreases, the length of unsteady phase becomes shorter, and the steady phase cloud becomes larger, making it impractical to distinguish these two phases from each other. Therefore,

application of phase diagram for automatically detecting steady state convergence is not possible when the number of particles is not large enough.

Figure 4 represents the distribution of samples from the statistical perspective for a system with different numbers of particles (N_i) per cell. From statistics, when the sampling procedure is free from errors and the sample size is large enough, the distribution function of samples is expected to be Gaussian. However, Fig. 4 shows that, even for a system with as many as 100,000 samples, which is enough in most problems, the distribution of samples is not Gaussian. In particular, the distribution functions of samples in the momentum (\mathbf{P}_i) and energy (ε_i) are shifted to the right as a result of inclusion of unsteady samples in the calculation—the type II error. To neutralize the effect of the type II error, more number of samples from the steady state will be needed. This is achieved in Fig. 4 (b) where the flow is sampled 10,000,000 times. The effect of the type II error on accuracy of the sampling procedure would be more critical in a system with a higher number of particles as can be seen in Fig. 4 (c).

The distribution function of the DSMC sampling indicates that the influence of the type II error on sampling accuracy becomes more important as the number of particles increases. In addition, the length of the unsteady transition phase becomes longer as the number of particles increases. This is caused by increment of the solution resolution and reduction of the statistical uncertainty. However, though the analysis of the distribution of samples provides much useful information regarding the sampling procedure, it cannot be used as a convergence criterion for steady state detection. This is because the generation of sample distribution functions is computationally expensive and the determination of deviation of a sample distribution from the Gaussian distribution is complicated. It is also due to the fact that the accuracy of the analysis of the distribution of samples is highly dependent on the number of particles. Hence, a new method to detect the onset of steady state by considering both the statistical characteristics of sampling and physical aspects of the problem is required.

2.2. Probabilistic automatic reset sampling (PARS)

This paper presents a new automatic method for detecting the onset of the steady state on the basis of the statistical features of the gas particles. We aim to develop a self-starter algorithm with no dependency on flow geometry and irrespective of the number of simulated particles. The method is then expected to be capable of detecting the onset of the steady state even when the number of simulated particles remains constant. In order to achieve a probabilistic automatic reset sampling (or an automatic steady state detection), two distinctive approaches—global and local—are developed.

2.2.1 Global steady state detection

Global steady state detection (or reset sampling) is based on the observation that collisional invariants—the number of particles, momentum, and kinetic energy—are continuously changing in the unsteady phase. In the unsteady phase, even though the number of particles remains constant, other components of collisional invariants may still be evolving. This behavior is clearly demonstrated in Fig. 5 of the test case of the 1-D Couette flow simulation. As explained in Fig. 2, higher order phase portraits provide more useful information for steady state detection. Therefore, if we consider the second and third components of collisional invariants (momentum and kinetic energy), instead of the first component (number of particles), which is the case of Bird's steady state detection method [21], a clear identification of the steady state is possible. In this (global) steady state detection method, three independent criteria corresponding to each of three collisional invariants can be derived. It is clear in Figs. 2 and 5 that variations of the number of particles, momentum and kinetic energy are bounded (with oscillations around the mean value) in the steady phase, while they are large in the unsteady phase. This statistical feature plays an essential role in the new automatic steady state detection method.

In the PARS algorithm, two set of time intervals, corresponding to a large number of simulation time steps, are needed in order to decide whether restarting or continuing the sample process. As a consequence, two set of variables are needed to be employed through the decision process. The first set of information is the reference value that can be obtained from the solution itself, and is being updated every time when the sampling process is restarted. The other set of information is the expected mean value of the particle number, momentum, and energy. In the PARS algorithm, the expected mean value, rather than instantaneous value, is used to calculate the desired properties. This method is adopted in order to avoid an inaccurate judgment about steady state and to take into account the statistical fluctuation of the properties in steady state. Hence, it is necessary to define a time interval period after the mean value is calculated. The value of the time interval should be chosen carefully, since the application of mean value throughout time interval may result in the false detection of the start point of the steady state, if the start point is within the time interval. Time intervals in the range of 100 to 1000 time steps are found appropriate for most DSMC applications.

Figure 6 describes the implementation of the new (global) steady state convergence detection algorithm in the existing DSMC method. First, the value of the total number of particles, momentum and kinetic energy are calculated based on the initial information. These values are then set as initial reference quantities. Next, the move and collision steps in the DSMC proceed. In contrast to the conventional DSMC algorithm, a steady state convergence detection module is called before sampling. In the convergence detection module, the collisional invariants are summed for all the particles in the simulation domain. The next step of the global PARS algorithm involves the calculation of expected averages and variance of collisional invariants according to the following relations [24, 25],

$$\langle N \rangle = \frac{\sum_{i=1}^{N_s} \theta_i}{N_s}, \ \langle P \rangle = \frac{\sum_{i=1}^{N_s} \upsilon_i}{N_s}, \ \langle E \rangle = \frac{\sum_{i=1}^{N_s} \varepsilon_i}{N_s}, \tag{1}$$

$$Var(N) = Var(\theta_{i}) = Var(N_{i}),$$

$$Var(P) = Var(\upsilon_{i}) = Var\left(\sum_{j=1}^{N_{i}} m \upsilon_{j}\right),$$

$$Var(E) = Var(\varepsilon_{i}) = Var\left(\sum_{j=1}^{N_{i}} m \upsilon_{j}^{2} \right).$$
(2)

Here, θ_i , v_i , ε_i are the sum of the number of particles, momentum (*mv*), and energy (*mv*²) with *m* and *v* being the molecular mass and velocity. N_s is the total number of samples during the simulation. $Var(\bullet)$ and $\langle \bullet \rangle$ are the variance and the expected mean average of the variable, respectively. N_i is the total number of particles at present sample step.

One way to determine the standard deviation, sample means, and confidence intervals is to calculate them theoretically. Assuming a simple mono-component gas in a sample cell with constant volume [26], it is possible to define the theoretical variance and standard deviation of the collisional invariants. By considering a simple ideal gas, represented within the grand canonical ensemble, the

probability of finding the system with N_i particles can be determined by the Poisson distribution [1, 27]. In most situations, a Poisson distribution function to model the unidirectional flux of particles crossing a plane is found fairly accurate approximation [23, 28, 29].

With the assumption of the Poisson distribution function, the variance, σ_N^2 , and standard deviation of the particles, σ_N , can be calculated based on the statistics and probability theory [25],

$$\sigma_N^2 = \langle N \rangle = N_{ref}, \ \sigma_N = \sqrt{N_{ref}} \ . \tag{3}$$

After defining the particle statistical properties, it is necessary to find the particle velocity and translational energy. The speed of the particles is governed by the Maxwell-Boltzmann distribution at equilibrium [1, 27, 30, 31] with the statistics defined as below,

$$f(v) = \left(\frac{m}{2\pi k_B T}\right)^{3/2} \left(4\pi v^2\right) e^{-\frac{mv^2}{2k_B T}},$$
(4)

$$\langle v \rangle = \int_{-\infty}^{+\infty} v f(v) dv = \sqrt{\frac{8k_B T}{\pi m}} \quad \text{and} \quad \langle v^2 \rangle = \int_{-\infty}^{+\infty} v^2 f(v) dv = \frac{3k_B T}{m},$$
 (5)

$$\sigma_{\nu} = \sqrt{\langle \nu^2 \rangle - \langle \nu \rangle^2} = \sqrt{\frac{3\pi - 8}{\pi} \frac{k_B T}{m}} = \sqrt{\frac{3\pi - 8}{8}} \langle \nu \rangle, \qquad (6)$$

where *m*, *v*, *T*, and k_B represent the particle mass, velocity, Boltzmann constant and bulk temperature, respectively. The f(v) is the Maxwell-Boltzmann distribution function, and $\langle v \rangle$, $\langle v^2 \rangle$ are the first and second moment of the velocity distribution function, respectively. The standard deviation of the particle velocity distribution function is denoted by σ_v . The translational energy distribution, f(E), can also be derived from the particle velocity distribution function as follows,

$$E = \frac{mv^2}{2} \quad \text{and} \quad dE = mvdv \tag{7}$$

$$f(v)dv = f(E)dE , \qquad (8)$$

$$f(E) = f(v)\frac{dv}{dE} = \left(\frac{4\pi v}{m}\right) \left(\frac{m}{2\pi k_B T}\right)^{3/2} e^{-\frac{mv^2}{2k_B T}}.$$
(9)

By defining $\beta = 1/k_B T$, the kinetic energy distribution can be written as a chi-squared distribution function with three degrees of freedom [25],

$$f(E) = 2\beta \frac{1}{\sqrt{2\pi}} \left(2\beta E\right)^{1/2} e^{-2\beta E} \,. \tag{10}$$

Further, the expected average $\langle E \rangle$, the standard deviation, σ_E , and the variance of the energy distribution function, σ_E^2 , can be calculated as below by setting ϕ (i.e., a scale parameter) equal to $1/2\beta$,

$$\langle E \rangle = 3\phi = \frac{3k_BT}{2}, \ \sigma_E^2 = 6\phi^2 = \frac{3}{2}(k_BT)^2, \ \sigma_E = \sqrt{\frac{6}{4\beta^2}} = \sqrt{\frac{2}{3}}\langle E \rangle.$$
 (11)

The standard deviation of the particle number σ_N , momentum σ_P , and kinetic energy σ_E can be calculated by considering (2), (3), (6), and (11),

$$\sigma_{N} = \left\langle N_{i}^{2} \right\rangle - \left\langle N_{i} \right\rangle^{2} = \sqrt{N_{i}},$$

$$\sigma_{P} = \left\langle P^{2} \right\rangle - \left\langle P \right\rangle^{2} = \sum_{j,k=1}^{N_{i}} Cov(mv|_{j},mv|_{k}) = \sqrt{\frac{3\pi - 8}{8}} \langle v \rangle,$$

$$\sigma_{E} = \left\langle E^{2} \right\rangle - \left\langle E \right\rangle^{2} = \sum_{j,k=1}^{N_{i}} Cov(\frac{mv^{2}}{2}|_{j},\frac{mv^{2}}{2}|_{k}) = \sqrt{\frac{2}{3}} \langle E \rangle.$$
(12)

Here $Cov(\bullet, \bullet)$ denotes the covariance of two random variables. The standard deviation of collisional invariants can also be calculated either empirically or theoretically in order to utilize in the calculation of the confidence intervals;

$$\sigma_{N} = \sqrt{N_{i}}, \ \sigma_{P} = \sqrt{\frac{1}{N_{i}} \left(\sum_{j=1}^{N_{i}} \left(mv_{j} - \langle v_{i} \rangle \right)^{2} \right)}, \ \sigma_{E} = \sqrt{\frac{1}{N_{i}} \left(\sum_{j=1}^{N_{i}} \left(\frac{mv_{j}^{2}}{2} - \langle \varepsilon_{i} \rangle \right)^{2} \right)}.$$
(13)

The next phase after computation of the expected averages and variance involves defining a restarting criterion based on the confidence intervals. The simplest way to define the reset criterion is to consider the variations of the number of particles, momentum, and energy with respect to the expected mean values as follows,

$$\left|N_{i}-N_{ref}\right| \leq Z_{\alpha_{1}}\sigma_{N_{ref}}, \left|P_{i}-P_{ref}\right| \leq Z_{\alpha_{2}}\sigma_{P_{ref}}, \left|E_{i}-E_{ref}\right| \leq Z_{\alpha_{3}}\sigma_{E_{ref}}.$$
(14)

Here, $\sigma_{N_{ref}}$, $\sigma_{P_{ref}}$, $\sigma_{E_{ref}}$ represent the standard deviation of the number of particles, momentum, and translational energy. The Z_a stands for z-score (real constant) of a Gaussian distribution function which can be adjusted to make the tradeoff between accuracy and efficiency. If the variations of the number of particles, momentum, and energy are higher than pre-defined ranges, the sampling process should be reset. These criteria are conceptually simple and easy to code, and, at the same time, they allow a broad range of variations for the momentum and energy. However, they may not be practical in dealing with the local unsteady regions where the variations of momentum and energy in the system are small. Moreover, the momentum and energy do not necessarily follow a Gaussian distribution function, thus selecting the values of Z_{α_2} and Z_{α_3} is not an easy task. In this regard, another reset criterion, based on *significance test* with the assumption of the negligible variations of particles, momentum, and energy variance at each time step, can be introduced. Here, rather than the absolute value of variance, a relative standard variation (coefficient of variation) of collisional invariants, \Re , is applied in order to integrate all three reset criteria into one, making it more sensitive to the momentum and energy variations during unsteady phase:

$$\Re = a_{reset} \frac{1}{\sqrt{N_i}} \min\left(\frac{\sigma_{E_i}}{\langle E_i \rangle}, \frac{\sigma_{P_i}}{\langle P_i \rangle}\right),\tag{15}$$

where a_{reset} is a constant value (close to 1) that can be adjusted by user to make a trade-off between accuracy and efficiency. It is recommended to define value of a_{reset} for the particle, momentum and energy criteria individually. Therefore, the sampling process may not be reset due to small statistical fluctuations in the particle number, momentum and energy. The recommended range of this parameter is $1.0 \le a_{reset} \le 4.0$ for most of the applications. The larger values lead to more speed up and less accuracy. The abovementioned relations are reduced to the following simple criterion:

$$1.0 - \Re < \max(\mathbb{R}_{Particle}, \mathbb{R}_{Momentum}, \mathbb{R}_{Energy}), \quad \min(\mathbb{R}_{Particle}, \mathbb{R}_{Momentum}, \mathbb{R}_{Energy}) < 1.0 + \Re. \quad (16)$$

In this expression in a range from $1.0 - \Re$ to $1.0 + \Re$, the parameters \mathbb{R} are defined as the ratio of collision invariants to their corresponding reference values:

$$\mathbb{R}_{Particle} = \frac{\langle N \rangle}{N_{ref}}, \ \mathbb{R}_{Momentum} = \frac{\langle P \rangle}{P_{ref}}, \ \mathbb{R}_{Energy} = \frac{\langle E \rangle}{E_{ref}}.$$
(17)

Here N_{ref} P_{ref} and E_{ref} stand for the reference number of particles, momentum, and translational energy. Notice that the ratios of the averaged parameter to reference value are utilized in deciding whether or not the samples should be reset. The use of an averaged parameter, instead of each step values, in the nominator of the ratios makes it possible to avoid an unnecessary sample reset due to statistical fluctuations.

2.2.2. Local steady state detection

The sample size in the DSMC simulation is determined by the number of particles (*N*) and sampling times performed for a cell (N_{SPC}). Since the cell is considered an independent statistical population in the DSMC, macroscopic properties of each cell can be calculated based on the sample information of the cell only, independent of other cells. Thus, each cell can have its own sample size, making it possible to develop the concept of *local* sampling with a variable sample size. The variable sample size can be achieved in two ways: one, by using a different number of particles (in each cell) at each time step, and another by using a different number of time steps that each cell is sampled. In the conventional DSMC algorithms, the local sampling was possible only through the former, thereby completely overlooking the possibility of achieving higher computational efficiency through the latter. In order to fully utilize the latter type of local sampling, it is necessary to accurately detect steady state at each cell and then perform sampling for that cell individually. Figure 7 summarizes the concept of local reset sampling in comparison with the conventional sampling procedure and explains schematically how a significant saving in overall computational cost can be achieved by adopting the local reset sampling.

Local steady state detection (or reset sampling) is based on a local variation of particles in each cell and there is no need of using the velocity and energy distribution functions; therefore sampling would start at each cell independently. Particles in the cell are assumed to follow the Poisson distribution function. This is the only assumption in the development of the local PARS method. First,

a time period corresponding to the number of samples, \wp , is defined to eliminate statistical fluctuations in the evaluation of the steady state at a cell. Therefore the algorithm can deal with the mean of the number of particles in a specific time interval rather than the instantaneous number of particles at each time step. The next step is to define the confidence interval of valid fluctuation of the particles in each cell. According to the Cramer-Rao lower band and minimum variance unbiased estimation [12], the sum of the mean of samples is a sufficient and complete quantity for describing the mean of the Poisson distribution function. Hence, the mean of population (\overline{N}) can be approximated using the maximum likelihood estimate,

$$\overline{N} = \frac{\sum_{i=1}^{\wp} N_i}{\wp},\tag{18}$$

and the corresponding Poisson distribution may also be expressed as

$$P(R) = \frac{(\overline{N})^R e^{-\overline{N}}}{\wp!},\tag{19}$$

where *R* represents the power of the mean of the distribution function, and \wp stands for the number of trials or samples. The variance and standard deviation of the mean of the population distribution can also be calculated as follows:

$$\sigma_N^2 = \overline{N}, \ \sigma_N = \sqrt{\overline{N}}.$$
 (20)

In order to find the trigger point for restarting sample process in the PARS algorithm, it is essential to find a relation between the variations of the mean of the particles at two different time periods, and to calculate the rest margin. Therefore, it is necessary to obtain the confidence interval of the mean of the Poisson distribution function by employing the relation between the chi-squared and cumulative Poisson distribution functions [32, 33]. The following confidence interval provides the allowable level of particle functions in the steady state.

$$\left[0.5\chi^{2}(\frac{\alpha}{2},2n),0.5\chi^{2}(1-\frac{\alpha}{2},n+2)\right].$$
(21)

Here $\chi^2(a,b)$ is a quantile function corresponding to a lower tail area (*a*) of the chi-squared distribution with *b* degrees of freedom. The *a* appearing in the first parameter of the chi-squared distribution is directly related to the confidence level of accuracy, 1-*a*, specified by user.

Nonetheless, the calculation of a desired confidence interval range using (21) is rather complicated [34, 35]. An alternate approximate formula may be derived by noting the relationship between the chi-squared and gamma distribution functions,

$$[F^{-1}(\alpha_{2}';n,1),F^{-1}(1-\alpha_{2}';n+1,1)],$$
(22)

where $F^{-1}(a;b,1)$ is a quantile function of gamma with the shape factor of *b*, scale factor of 1, and lower tail area of *a*. This formula, however, turns out to be still computationally expensive. Thus, a more simplified formula of the confidence level can be obtained by the Wilson-Hilferty transformation as follows [36],

$$\left[n\left(1-\frac{1}{9n}-\frac{Z_{\frac{n}{2}}}{3\sqrt{n}}\right)^{3},(n+1)\left(1-\frac{1}{9(n+1)}+\frac{Z_{\frac{n}{2}}}{3\sqrt{n+1}}\right)^{3}\right].$$
(23)

Here $Z_{\alpha/2}$ is the standard normal deviation with upper tail area $\alpha/2$ and *n* represents the mean of number of particles in the cell ($n = \overline{N_a}$). Finally, based on this formula, the lower and upper limits of the confidence level are defined as follows,

Lower Limit =
$$\left[n \left(1 - \frac{1}{9n} - \frac{Z_{\frac{\alpha}{2}}}{3\sqrt{n}} \right)^3 \right],$$
 (24)
Upper Limit = $\left[(n+1) \left(1 - \frac{1}{9(n+1)} + \frac{Z_{\frac{\alpha}{2}}}{3\sqrt{n+1}} \right)^3 \right].$

In the local steady state detection algorithm, two successive time periods (each containing multiple DSMC steps) are used to investigate the onset of steady state. The mean of the number of particles in a cell, \overline{N}_a , is calculated at the end of the first period and is set as the reference parameter, N_{ref} . Based on this reference value, the upper and lower limits of the confidence level are also calculated by using (24). Then, another mean number of particles in the cell, \overline{N}_b , is calculated at the end of the second period, and is checked to determine whether or not it is inside the confidence interval. If the \overline{N}_b is inside the confidence interval, it implies that the local cell already reached the

steady state. Otherwise, sampling should be reset and the \overline{N}_b is set as the new reference value. Consequently, the upper and lower limits of confidence levels are updated based on new reference value. The process is repeated until the \overline{N}_b is inside the interval. This completes the algorithm of the local steady state detection.

3. Results and Discussions

In order to confirm the capability of the new steady state detection method base on the automatic reset sampling, a benchmark problem, the 1-D shear-driven Couette flow is investigated. The Couette flow is defined as the flow trapped between two infinite, parallel, flat plates at $x = \pm H$ driven by the shear motion of one of the plates with a Mach number (M=1 in the present case), while the temperature of the walls is maintained constant (293 K in the present case). It is also assumed that the fluid moves in the y-direction only. The Knudsen number based on the characteristics length, i.e. the gap between plates, is set 0.25 in the test cases. The study on the simple 1-D Couette flow allows us to avoid the geometrical complexity and focus solely on the sampling procedure. Moreover, one of the main limitations of the existing methods is their inability to handle problems with a constant number of particles. Therefore, the 1-D Couette flow is ideal for evaluating the performance of various (existing and new) methods.

Figure 8 shows the typical variations of various macroscopic properties at the central cell of the domain with respect to the number of sample steps. An important property, the total kinetic energy of the system, is also included for detailed comparison. In this no-reset simulation, all the macroscopic properties approach asymptotic values as the total kinetic energy level is stabilized. This is due to the fact that all the properties, including the kinetic energy, are directly dependent on the molecular velocity.

Figure 9 shows the variations of various macroscopic properties with respect to the number of sample steps with and without applying the global PARS so that the effects of the global PARS may be clearly identified. When the PARS is active, the macroscopic variables in general tend to approach the final converged values with fewer samples taken in the simulation and, at the same time, less fluctuation around the true average values. However, the final solution will become identical for both

methods, and unsteady sampling effects will vanish, since very large number of samples is taken during the simulation. In particular, this feature is apparent for the properties of the density, pressure, temperature, *v*-velocity and shear stress. On the other hand, the *x*, *z* components of the particle velocity, and the heat flux—which is the function of all velocity components, including *x*, *z* components—do not respond as fast as the other macroscopic properties do. This is due to the peculiarity of the present 1-D Couette flow problem in which the macroscopic properties vary in the *y*-coordinate only. In passing, it should be noted that fluctuations appearing near the converged values are generic in the DSMC and they will attenuate to negligible level once enough samples are taken.

Figure 10 illustrates the effects of the number of particles on the total kinetic energy when no reset sampling is applied. First, since the total kinetic energy is the sum of the energy of all particles of the system, the system with different numbers of particles will possess different energy levels, as confirmed in Fig. 10 (a). Second, it can be observed that the standard deviation of the energy depends highly on the number of particles and the total kinetic energy exhibits highly oscillatory behavior, in particular, for the system with a small number of particles. Therefore, the instantaneous level of the total kinetic energy may not be suitable for detecting the steady state. Instead, the ratio of the average kinetic energy would be preferable, the behavior of which is shown in Fig. 10 (b). In contrast to Fig. 10 (a) of the total kinetic energy, all the ratios for different numbers of particles collapse on a single curve.

It is expected that the energy ratio should approach unity if both the average energy and the reference value belong to the steady state. As shown clearly in Fig. 10 (b), however, the energy ratio does not converge to unity when no reset sampling is applied, irrespective of the number of particles per cell. This is exactly why the reset sampling is required for accurately detecting the steady state. Figure 11 compares the ratio of the average kinetic energy with and without applying the reset sampling. It demonstrates that the energy ratio converges nicely to the expected value after some oscillations around unity. Note also that each peak represents the point at which the reset sampling is applied.

Figure 12 shows the variations of normalized momentum and kinetic energy with respect to the number of sample steps for two different numbers of particles. Each symbol represents the reset iteration point at which the global PARS is applied. It can be seen that the system with smaller number of particles reaches the steady state earlier. This can be explained as follows: in case of simulations with smaller number of particles, there may be only few particles in unsteady regions, leading to premature satisfaction of the steady state condition. On the other hand, the statistical properties in simulations with smaller sample size, i.e., simulations with less than 20 particles per cell, are contaminated with the large amount of statistical uncertainty, leading again to premature satisfaction of the criteria.

In order to analyze the concept of local steady state detection in detail, the 2-D high speed flow past a vertical wall is investigated with and without applying the local PARS. This flow geometry is chosen because, while it is simple, it is nonetheless able to produce fairly rich flow features, such as gaseous compression, expansion, recirculation, and also low density regions behind the wall. Indeed, the vertical wall flow was suitable for verifying the concept of local PARS. Figure 13 compares the simulation results of conventional Bird's and new local PARS algorithms at the Mach number 2.0 and the Knudsen number 0.04. The free-stream number density and temperature are set to be 1.0×10^{19} and 300 K, respectively. The time step is chosen to be sufficiently smaller than the mean collision time in order to satisfy the time step constraint. Figure 13 (a) is obtained for the system with maximum 340 particles per cell by Bird's conventional method in which the sampling procedure starts after all the sample cells converge to the steady state. On the other hand, Figs. 13 (b) and (c) are obtained by the local reset sampling method for the system with maximum 340 and 30 particles per cell, respectively.

The comparison of three Mach number contours in Fig. 13 indicates that the macroscopic properties are not degraded even with fewer iteration, in the case of the local steady stated detection method. The comparison of cases with different numbers of particles (Figs. 13 (b) and (c)) reveals that the new method works successfully even for the system with few particles as small as 20. It can also be observed that the efficiency of the model decreases with a decreasing number of particles per cell.

Moreover, due to a large confidence interval for the system with a small number of particles, the unsteady regions are not fully resolved. However, considering the fact that at least 20 particles per cell is necessary for a proper DSMC simulation, there are, in principle, no limitations for the method. In addition, it can be seen in the contours of N_{SPC} (i.e., the number of samples at each cell) that the steady state is reached upstream from the vertical wall earlier than in other regions, while downstream behind the vertical wall remains highly unsteady with a smaller number of samples taken. Note also that the contours of the samples in the local reset sampling method not only detect local unsteady regions but also help in determining the ideal size of the computational domain, since the flow reaches the steady state from the far-field earlier than other regions.

Figure 14 shows the steady state convergence trend of the local reset sampling for different numbers of particles per cell. In this figure, the number of cells that is still in the unsteady phase is reported before the steady state is reached. The number of unsteady cells decreases as the simulation proceeds. While the majority of cells reach the steady state in less than 2,000 iterations, few cells (less than 300 hundreds) need much more time to converge to the steady state. In addition, the system with fewer particles per cell converges to the steady state faster, which is consistent with the result obtained from the global steady state detection method. This is again due to the increased confidence interval in the system with fewer particles.

Finally, Fig. 15 compares the computational cost of conventional, global and local PARS steady state detection methods in the 2-D high speed flow past a vertical wall. The relative CPU time is obtained by normalizing the computational time based on the non-automatic conventional steady state convergence method. For this calculation, the same number of iteration, 10,000, is used for N_{SPC} in the global and local PARS methods and the number of samples in the conventional method. In the case of the local steady state detection method, the sampling procedure starts when 99% of the sampled cells converge to the steady state. The results show that application of the global steady state detection method leads to 6.4%, 6% and 14% gain for systems with a maximum of 340, 170, and 30 particles per cell, respectively. These gains are due to the automatic detection of steady state convergence over the predefined (non-automatic) start of sampling.

On the other hand, application of the local steady state detection method leads to the substantial gain of 30%, 32% and 36%, respectively. These gains are possible by the combined effects of the automatic detection and local reset sampling. Note also that the local reset sampling—easily implementable on parallel platforms—outperforms the automatic detection of steady state convergence in computational savings, 36% versus 14%, in the case of a maximum of 30 particles per cell. This comparison states that the global steady state detection method outperforms the traditional steady state method; however, it is not as efficient as the local steady state detection method. This is because, in the global detection method, the sampling process is postponed until the whole domain reaches the steady state condition, whereas, in the local detection method, the sampling process starts immediately for the cells that satisfy the steady state condition.

4. Conclusions

Owing to significant efforts in algorithm development and the rapid progress in computing power in past decades, the DSMC is now established as a primary workhorse to computationally solve the kinetic Boltzmann equation and is routinely being applied to various flow problems of scientific and technological interests. However, as is common when the laboratory level research of computational models enters the real world, there are several hurdles for the DSMC practitioners to overcome before they can use it for complicated applications in an effective way.

One such example is the case of how to determine precisely when to initiate (or stop) the sampling for obtaining a steady state solution. In the case of lacking fully automatic robust methods to handle such problems, effective and uninterrupted use of the DSMC is achievable after many trial and errors for a considerable period of time. Moreover, as applications of the DSMC diversify, there is additional need to handle challenging boundary conditions such as adjusting the pressure at a pressure-outlet zone of an air intake in order to meet the desired mass flow rate, demanding the capability of the user-input-free black-box style like recent commercial CFD codes.

The main goal of the present study was to propose an ingenious method based purely on the statistical features of the gas particles as a step toward developing a robust fully automatic steady state detection method. The key concept was built upon automatic restarting sampling to minimize the type

II error caused by incorrect acceptance of the samples that do not belong to the steady state. In order to investigate in detail the effect of incorrect inclusion of samples, phase portraits of the sampling estimators and distribution of samples were considered. The higher order phase portraits were found to provide the best illustration of the transition phase. However, though the phase portrait analysis provided much useful information regarding the sampling procedure, we judged it not applicable to deriving a robust convergence criterion for steady state detection, primarily due to the fact that the accuracy of the phase portrait analysis is highly dependent on the number of particles.

On the basis of the observation that the Boltzmann collisional invariants are continuously changing in the unsteady phase, new fully-automatic global and local steady state detection methods were developed. The global steady state method is based on a relative standard variation of collisional invariants, enabling it to be more sensitive to the momentum and energy variations during the unsteady phase. To verify the capability of the new method, the 1-D shear-driven Couette flow was extensively investigated. It was shown that, with the new method, the macroscopic variables tend to approach the final converged values with fewer samples taken in the simulation.

Moreover, by recognizing the fact that the cell in the DSMC can be considered an independent statistical population, the local steady state method based on local variations in the distribution function of particles at each cell was developed. A simple formula based on the Wilson-Hilferty transformation was derived to provide information on the confidence interval for the reset sampling procedure. To verify the concept of local steady state detection, the 2-D high speed flow past a vertical wall was also investigated. It was shown that the local method works successfully even for the system with few particles as small as 2. Even though the application of local PARS algorithm was reported only for external supersonic flow in the present study for the sake of simplicity, it is applicable for internal and external low speed flows as well (for example, lid-driven cavity flows). Because the present method is based on statistical variation of the particles in a sample cell, not macroscopic field properties, the local PARS algorithm is expected to perform properly for all speed regimes. Nonetheless, the efficiency of the method in low speed flows may depend highly on the number of particles per cell and speed of information transition in the computational domain. In case of lid-driven cavity flow, the local steady state detection method turned out to lead to the gain of 8.5%

compared to the traditional steady state counterpart. The decrease of the gain may be explained by the fact that, due to higher statistical fluctuations present in low speed flow conditions, larger number of simulation particles is required for accurately locating the unsteady regions.

In summary, owing to the combined effects of the automatic detection and local reset sampling, the local steady state detection method was found to yield a substantial gain of 30-36% for the problem studied. Interestingly, the local reset feature was found to outperform the automatic detection feature in overall computational savings. This suggests that further potential saving in computational cost may be made by refining the local reset sampling concept.

ACKNOWLEDGMENTS

This work was supported by the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (Basic Science Research Program NRF 2015-M1A3A3A02-01062), South Korea.

REFERENCES

[1] G.A. Bird, Molecular gas dynamics and the direct simulation Monte Carlo of gas flows, Oxford Engineering Science Series, 1994.

[2] E. Oran, C. Oh, B. Cybyk, Direct simulation Monte Carlo: Recent advances and applications, Annual Review of Fluid Mechanics, 30 (1998) 403-441.

[3] W. Wagner, A convergence proof for Bird's direct simulation Monte Carlo method for the Boltzmann equation, Journal of Statistical Physics, 66 (1992) 1011-1044.

[4] A. Karchani, R.S. Myong, Convergence analysis of the direct simulation Monte Carlo based on the physical laws of conservation, Computers & Fluids, 115 (2015) 98-114.

[5] J.S. Wu, K.C. Tseng, Analysis of micro-scale gas flows with pressure boundaries using direct simulation Monte Carlo method, Computers & Fluids, 30 (2001) 711-735.

[6] Q. Sun, I.D. Boyd, A direct simulation method for subsonic, microscale gas flows, Journal of Computational Physics, 179 (2002) 400-425.

[7] J.S. Wu, K.C. Tseng, T.J. Yang, Parallel implementation of DSMC using unstructured mesh, International Journal of Computational Fluid Dynamics, 17 (2003) 405-422.

[8] M. Darbandi, A. Karchani, G. Schneider, The study of rarefied gas flow through microfilters with different openings using MONIR-DSMC, in: ICNMM2011, ASME, Alberta, Canada, (2011) 215-221.

[9] M.Y. Plotnikov, E. Shkarupa, Theoretical and numerical analysis of approaches to evaluation of statistical error of the DSMC method, Computers & Fluids, 105 (2014) 251-261.

[10] M.Y. Plotnikov, E. Shkarupa, Estimation of the statistical error of the direct simulation Monte Carlo method, Computational Mathematics and Mathematical Physics, 50 (2010) 335-344.

[11] A. Karchani, O. Ejtehadi, R.S. Myong, A steady-state convergence detection method for Monte Carlo simulation, AIP Conference Proceedings, 1628 (2014) 313-317.

[12] H. Cramér, Mathematical methods of statistics, Princeton University Pres, 1999.

[13] G. Chen, I.D. Boyd, Statistical error analysis for the direct simulation Monte Carlo technique, Journal of Computational Physics, 126 (1996) 434-448.

[14] T.E. Schwartzentruber, L.C. Scalabrin, I.D. Boyd, A modular particle–continuum numerical method for hypersonic non-equilibrium gas flows, Journal of Computational Physics, 225 (2007) 1159-1174.

[15] M.A. Gallis, J. Torczynski, D. Rader, G.A. Bird, Convergence behavior of a new DSMC algorithm, Journal of Computational Physics, 228 (2009) 4532-4548.

[16] Q. Sun, J. Fan, I.D. Boyd, Improved sampling techniques for the direct simulation Monte Carlo method, Computers & Fluids, 38 (2009) 475-479.

[17] N.G. Hadjiconstantinou, A.L. Garcia, M.Z. Bazant, G. He, Statistical error in particle simulations of hydrodynamic phenomena, Journal of Computational Physics, 187 (2003) 274-297.

[18] H.M. Cave, K.C. Tseng, J.S. Wu, M.C. Jermy, J.C. Huang, S.P. Krumdieck, Implementation of unsteady sampling procedures for the parallel direct simulation Monte Carlo method, Journal of Computational Physics, 227 (2008) 6249-6271.

[19] G.A. Bird, Molecular gas dynamics, Oxford Engineering Science Series, 1976.

[20] M.A. Rieffel, A method for estimating the computational requirements of DSMC simulations, Journal of Computational Physics, 149 (1999) 95-113.

[21] G.A. Bird, Sophisticated DSMC, Notes from DSMC07, Santa Fe, 2007.

[22] G.A. Bird, The DSMC method, CreateSpace Independent Publishing Platform, 2013.

[23] J.M. Burt, I.D. Boyd, Convergence detection in direct simulation Monte Carlo calculations for steady state flows, Communications in Computational Physics, 10 (2011) 807.

[24] Y.A. Shreider, Method of statistical testing: Monte Carlo method, Elsevier Publishing Company, 1964.

[25] E. Parzen, Modern probability theory and its applications, John Wiley & Sons, 1960.

[26] S.M. Ermakov, G.A. Mikhailov, Statistical modeling, Nauka, Moscow, 1982.

[27] W.G. Vincenti, C.H. Kruger, Introduction to physical gas dynamics, New York, Wiley, 1965.

[28] M.K. Ochi, Applied probability and stochastic processes in engineering and physical sciences, Wiley, New York, 1990.

[29] A. Kostinski, A. Jameson, On the spatial distribution of cloud particles, Journal of the atmospheric sciences, 57 (2000) 901-915.

[30] I. MacDonald, A relation between the Maxwell-Boltzmann and chi-squared distributions, Journal of Chemical Education, 63 (1986) 575.

[31] A.L. Garcia, B.J. Alder, Generation of the Chapman–Enskog distribution, Journal of Computational Physics, 140 (1998) 66-70.

[32] F. Garwood, Fiducial limits for the Poisson distribution, Biometrika, (1936) 437-442.

[33] N.L. Johnson, A.W. Kemp, S. Kotz, Univariate discrete distributions, John Wiley & Sons, 2005.

[34] D.C. Montgomery, G.C. Runger, N.F. Hubele, Engineering statistics, John Wiley & Sons, 2009.

[35] C. Forbes, M. Evans, N. Hastings, B. Peacock, Statistical distributions, John Wiley & Sons, 2011.

[36] N.E. Breslow, N.E. Day, Statistical methods in cancer research, International Agency for Research on Cancer, 1987.

Methods	Features
Conventional non-automatic method [1]	Simple conceptUser input necessaryProblem-dependency
First automatic method [21]	 Detection based on variation of the total number of particles in the domain Free from the user input Not applicable to local steady-state detection Difficulty in handling highly unsteady regions (i.e., recirculating flow)
Automatic method based on boundary properties [23]	 Detection based on maximum local time variation in outward-directed number flux along the boundaries of the simulation domain and Skellam distribution Free from the user input Not applicable to local steady state detection Large interval time step requirement
New automatic method – global	 Detection based on the statistical variation of collisional invariants and chi-square distributions of particle energy Free from the user input Applicable to problems with a constant number of particles in the domain Not applicable to local steady state detection
New automatic method – local	 Detection based on the local variation of particles in the cells Free from the user input Capable of detecting highly local unsteady regions Potential application in parallel codes Substantial computational saving

Table 1. A summary of steady-state detection methods for the DSMC method



Fig. 1. Computational errors, including the statistical error, in the DSMC simulation.



Fig. 2. Phase portrait of random variables in the DSMC without the reset sampling (2,000 particles per cell). Among 15 possible combinations of variables (C_x, C_y, C_z) and (C_x^2, C_y^2, C_z^2) , only 3 representative cases are plotted.



(a) 20 particles per cell





Fig. 3. Effect of the number of particles on the phase portrait diagram. The portrait (with enough number of particles) in (C_x^2, C_y^2) is suitable for detecting the steady-state.



Fig. 4. Distribution of samples from the statistical perspective for systems with different numbers of particles per cell and samples.



Fig. 5. Variations of collisional invariants in the DSMC simulation with a constant number of particles (the one-dimensional Couette flow problem); (a) total number of particles, (b) velocity, and (c) energy.



Fig. 6. Implementation of the new (global) automatic steady state detection algorithm in the existing DSMC method; (a) conventional DSMC algorithm, (b) additional algorithm for automatic detection.



Fig. 7. Explanation of the core concept of the local reset sampling and how a saving in overall computational cost can be achieved by adopting the local reset sampling in comparison with the conventional sampling procedure.



Fig. 8. Convergence behavior of macroscopic properties (by the broken lines) versus the number of sample steps. The total kinetic energy of system by the solid lines is also included for comparison.



Fig. 9. Comparison of convergence behavior of macroscopic properties with (by the broken lines) and without applying the global PARS (by the solid lines).







Fig. 10. The effect of the number of particles; (a) total kinetic energy; (b) ratio of the average kinetic energy.



Fig. 11. Comparison of the ratio of the average kinetic energy with and without applying the PARS ($N_{per cell}=20$).





(b) **Fig. 12.** Reset iteration number–at which the sampling procedure is reset–for different numbers of particles; (a) normalized momentum; (b) normalized energy.



Mach contour

Particles per cell contour

 N_{SPC} contour

Fig. 13. Verification of the variable sampling procedure (i.e., local sampling) in twodimensional DSMC problem for the system with different numbers of particles per cell; (a) no variable (conventional Bird) sampling (a maximum of 340 particles per cell); (b) variable sampling (a maximum of 340 particles per cell); (c) variable sampling (a maximum of 30 particles per cell). N_{SPC} denotes the number of samples at each cell.



Fig. 14. Steady state convergence trend for two-dimensional benchmark problem (with the local reset sampling).



Fig. 15. Comparison of conventional, global, and local steady state detection methods in terms of computational cost for two-dimensional benchmark problem.