# Statistical field estimators for multiscale simulations

Jacob Eapen,[1] Ju Li,[2] and Sidney Yip[1]

[1]*Department of Nuclear Science and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[2]*Department of Materials Science and Engineering, Ohio State University, Columbus, Ohio 43210, USA*
(Received 13 June 2005; published 30 November 2005)

We present a systematic approach for generating smooth and accurate fields from particle simulation data using the notions of statistical inference. As an extension to a parametric representation based on the maximum likelihood technique previously developed for velocity and temperature fields, a *nonparametric* estimator based on the principle of maximum entropy is proposed for particle density and stress fields. Both estimators are applied to represent molecular dynamics data on shear-driven flow in an enclosure which exhibits a high degree of nonlinear characteristics. We show that the present density estimator is a significant improvement over *ad hoc* bin averaging and is also free of systematic boundary artifacts that appear in the method of smoothing kernel estimates. Similarly, the velocity fields generated by the maximum likelihood estimator do not show any edge effects that can be erroneously interpreted as slip at the wall. For low Reynolds numbers, the velocity fields and streamlines generated by the present estimator are benchmarked against Newtonian continuum calculations. For shear velocities that are a significant fraction of the thermal speed, we observe a form of shear localization that is induced by the confining boundary.

## I. INTRODUCTION

An accurate field description is an essential ingredient in multiscale modeling strategies for coupling molecular simulation to continuum calculations [1–7]. A spatial buffer is typically prescribed for transferring data between the molecular and the continuum domains, with the former providing the interface or boundary conditions and the latter determining the mean field conditions for the molecular simulation. An alternative approach is to bypass the continuum equations and work exclusively with a series of numerical operations akin to the buffer transfers in coupled continuum-particle simulations [8–12]. The idea is to use short bursts of appropriately initialized microscopic simulations to estimate the evolution of the pertinent macroscopic field variables. This procedure relies on the continuity and smoothness of the macroscopic variables both in space and time, and the separation of time scales between the molecular and continuum dynamics. Regardless of the validity of a particular hypothesis, smooth and accurate field estimation is an intrinsic part of any multiscale methodology.

In this work we present a systematic approach to construct smooth and accurate fields from particle data using statistical inference techniques. A parametric method based on maximum likelihood inference is previously formulated to generate velocity and temperature fields [13] in situations where the particles in the system may be assumed to follow a local Maxwellian distribution. This method, however, cannot be applied to macroscopic variables such as the density and stress fields for which the form of the underlying distributions is not known *a priori*. Here we propose a nonparametric method in which the distribution is determined according to the principle of maximum entropy, with constraints in the form of moments that can be determined from the particle data. The maximum-entropy method identifies the distribution which is "maximally noncommittal with regard to missing information and that which agrees

with what is known, but expresses maximum uncertainty with respect to all other matters" [14]. To investigate the accuracy and smoothness of this estimator we carry out several comparisons of the density fields obtained in this way with those estimated by placing a smoothing kernel function over the data points, a characteristic procedure employed in smooth particle hydrodynamics (SPH) simulations [15].

The simplest way to bring out the mean-field behavior in particle simulations is through bin averages. Besides being rather arbitrary, this procedure also suffers from poor resolution [16,17] and statistical noise [18,19]. When the bin sizes are too small, the fields have a rough and jagged topology, and if the bins are too large, significant local information can get smeared out. The disadvantages of bin averaging are partly alleviated by the kernel method just mentioned. The field variable at any location is estimated by averaging the discrete data points over an arbitrarily chosen kernel [20,21]. Smoothness is ensured by the continuity and differentiability of the kernel function. While the kernel method is expected to be superior to bin averaging, it is not optimal in that the size (support) and shape (functional form) of the kernel remain arbitrary to some extent.

In comparing the maximum-entropy (ME) formulation with the kernel method, we first reconstruct a known function with nonlinear oscillations. We find the former gives a smooth estimate which captures the sharp gradients, while the latter, though also smooth, fails to resolve the larger gradients especially near the boundaries. Both methods are then applied to a molecular dynamics (MD) simulation of shear driven flow in an enclosure, a more complex problem than the commonly studied Coutte and Poiseuille flows. We find that the ME density estimate is a significant improvement over the bin average. Density deficiency near the boundaries, a common drawback of the kernel estimates, is conspicuously absent in the ME field.

We then examine the velocity fields given by the maximum likelihood (ML) and the kernel estimators. We again

find edge artifacts in the latter which left unidentified can lead to erroneous interpretation of velocity slip at the boundaries. For generating closed streamlines, we solve the generalized vorticity-streamfunction equation instead of directly integrating the velocity field. We find that this approach introduces additional smoothness and performs better than the direct integration scheme. Furthermore, we carry out continuum simulations using incompressible and compressible Navier-Stokes equations. For low Reynolds numbers, the ML fields are remarkably similar to those from the continuum simulations, indicating the Newtonian character of the MD flow field.

The plan of the paper is as follows. The theoretical basis and numerical implementation of the statistical field estimators are described in Sec. II with particular emphasis on the ME method. The method of kernel-based averages is also discussed for comparative studies. The model problem of shear or lid-driven flow in an enclosure is discussed in Sec. III along with the details of MD and continuum simulations. The field estimates are presented in Sec. IV and concluding remarks are made in Sec. V.

## II. STATISTICAL FIELD ESTIMATION

### A. Maximum entropy method

In statistical inference estimation, the field of interest is regarded as a probability distribution function. For the particle density and stress fields, the distribution is not known beforehand which necessitates the use of a nonparametric method. In this study, we employ the maximum entropy method with constraints in the form of moments. We adopt the univariate variational methodology of Mohammad-Djafari [22] and extend it to multivariate formalism. The basic problem in ME is to obtain a distribution which satisfies the imposed constraints while maximizing the uncertainty regarding the missing information, which in turn is measured by the information entropy. The desired distribution is written as

$$\hat{\rho}(\mathbf{x}) = \arg \max H[\rho], \qquad (1)$$

where $H[\rho]$ is the information entropy functional and $\rho(\mathbf{x})$ is the field of interest. For quantifying the entropy we use the definition of Shannon [23] which is given by

$$H[\rho] = - \int \rho(\mathbf{x}) \ln \rho(\mathbf{x}) d\mathbf{x}. \qquad (2)$$

The measure of information entropy is not unique; for example, other definitions such as Renyi's entropy [24] may be equally appropriate. The choice of (2) recognizes the role of thermodynamic entropy in statistical mechanics.

The maximization in (1) is to be carried out subject to certain constraints, which in the present problem represent the information from the particle data produced by molecular simulations. We will work with constraints in the form of expectation values or "generalized" moments of the distribution,

$$\int \varphi_{\mathbf{m}}(\mathbf{x}) \rho(\mathbf{x}) d\mathbf{x} = \mu_{\mathbf{m}}, \qquad (3)$$

where $\{\varphi_{\mathbf{m}}\}$ is a set of scalar functions such as polynomials or complex exponentials, $\mu$ represent the moments, and $\mathbf{m}$ is a $D$ dimensional vector with $D \leq 3$. Thus using polynomial functions $\{\mathbf{x}^{\mathbf{m}}\}$ as the moment-generating functions (here $\{\mathbf{x}^{\mathbf{m}}\}$ denotes a scalar function in $\mathbf{x}$ indexed by $\mathbf{m}$: $\mathbf{x}^{\mathbf{m}} \equiv x_1{}^{m_1} x_2{}^{m_2} \dots x_D{}^{m_D}$) gives the familiar geometric moments associated with the distribution which is given by

$$\int \mathbf{x}^{\mathbf{m}} \rho(\mathbf{x}) d\mathbf{x} = \mu_{\mathbf{m}}. \qquad (4)$$

Since the distribution is the quantity for which the estimator will provide, it may appear that the moments (3) are also unknown. However, the particle positions $\mathbf{X}_n$ are known from a molecular simulation and therefore one can formally write the particle density field at any point as

$$\rho(\mathbf{x}) = \frac{1}{N} \sum_{n=1}^{N} \delta(\mathbf{x} - \mathbf{X}_n), \qquad (5)$$

where $N$ is the total number of particles in the system. Inserting (5) into (3) we get

$$\mu_{\mathbf{m}} = \frac{1}{N} \sum_{n=1}^{N} \varphi_{\mathbf{m}}(\mathbf{X}_n), \qquad (6)$$

so the generalized moments of the density distribution can be expressed in terms of particle data. Similarly, the momentum distribution is

$$\mathbf{p}(\mathbf{x}) = \rho(\mathbf{x}) \mathbf{v}(\mathbf{x}) = \frac{1}{N} \sum_{n=1}^{N} V(\mathbf{X}_n) \delta(\mathbf{x} - \mathbf{X}_n), \qquad (7)$$

where $\mathbf{v}(\mathbf{x})$ is the velocity field, and $\mathbf{V}(\mathbf{X})$ are the particle velocities. The corresponding moments of the velocity field then become

$$\mu_{\mathbf{m}} = \frac{1}{N} \sum_{n=1}^{N} \mathbf{V}(\mathbf{X}_n) \varphi_{\mathbf{m}}(\mathbf{X}_n). \qquad (8)$$

For other distributions such as temperature and stresses similar results can be evaluated [25]. We will not need these expressions since in this work our interest lies only in the particle density field estimation.

Returning to (1), since the entropy functional is known to be concave [24], a unique maximum therefore exists for (2). Maximizing the entropy is equivalent to maximizing the Lagrangian

$$\mathcal{L} = H[\rho] + \sum_{\mathbf{m}} \int \lambda_{\mathbf{m}} \varphi_{\mathbf{m}}(\mathbf{x}) \rho(\mathbf{x}) d\mathbf{x}, \qquad (9)$$

where $\{\lambda_{\mathbf{m}}\}$ denote the Lagrange multipliers. The above expression can be rewritten as

$$\mathcal{L} = - \int \left\{ \rho(\mathbf{x})\ln \rho(\mathbf{x}) + \sum_{\mathbf{m}} \lambda_{\mathbf{m}} \varphi_{\mathbf{m}}(\mathbf{x})\rho(\mathbf{x}) \right\} d\mathbf{x}. \quad (10)$$

The desired density distribution is obtained by a functional differentiation with respect to $\rho(\mathbf{x})$ and equating to zero

$$\frac{\delta \mathcal{L}}{\delta \rho(\mathbf{x})} = 0 \quad (11)$$

from which one finds

$$\rho(\mathbf{x}) = \exp\left( -\sum_{\mathbf{m}} \lambda_{\mathbf{m}} \varphi_{\mathbf{m}}(\mathbf{x}) \right). \quad (12)$$

This is an exact expression for the distribution with the Lagrange multipliers $\{\lambda_{\mathbf{m}}\}$ as yet undetermined. The latter are given by the set of nonlinear equations formed by substituting (12) in (3)

$$\int \varphi_{\mathbf{m}}(\mathbf{x}) \exp\left( -\sum_{\mathbf{m}} \lambda_{\mathbf{m}} \varphi_{\mathbf{m}}(\mathbf{x}) \right) d\mathbf{x} = \mu_{\mathbf{m}}. \quad (13)$$

There are no closed from solutions for the Lagrange multipliers and hence, they are solved numerically by the Newton-Raphson method. We first expand the left side of (13), denoted as $G_{\mathbf{m}}(\lambda_{\mathbf{m}})$, in a Taylor series about trial values of the eigenvalue vector $\lambda_{\mathbf{m}}^0$, and neglect higher order derivatives to obtain a set of linear equations,

$$G_{\mathbf{m}}(\lambda_{\mathbf{m}}) = G_{\mathbf{m}}(\lambda_{\mathbf{m}}^0) + (\lambda_{\mathbf{m}} - \lambda_{\mathbf{m}}^0) \nabla G_{\mathbf{m}}(\lambda_{\mathbf{m}}) \quad (14)$$

which are then solved iteratively for $\lambda_{\mathbf{m}}$ until the relative change in the entropy ($H$) becomes less than a preassigned tolerance value. This method is second order accurate.

The moment-generating functions $\varphi_{\mathbf{m}}(\mathbf{x})$ can be viewed as the basis sets for $\ln \rho(\mathbf{x})$. We have used two basis sets, polynomials and complex exponentials, $\{\mathbf{x}^{\mathbf{m}}\}$ and $\{\exp(i\mathbf{m}\cdot\mathbf{x})\}$. Since the latter form an orthogonal basis set, they can be expected to have good numerical stability. For higher dimensional systems ($D > 1$), the basis set can in general be constructed as a product of univariate functions; for example, in two dimensions,

$$\varphi_{\mathbf{m}}(\mathbf{x}) = \varphi_{m_1}(x)\varphi_{m_2}(y). \quad (15)$$

We test the density field estimator using two kinds of data, the first is data sampled from a known distribution, while the second is produced in an MD simulation of shear-driven flow in a cavity, to be described in Sec. IV. In the first instance, we generate a data set of 15 000 points from the following distribution [27] by using the acceptance-rejection method [26]:

$$\rho(x) = C\left[ 1.5 + \sin(3\pi x) + \sin\left(\frac{15}{2}\pi x\right) \right], \quad (16)$$

where $C$ is a normalization constant. Then we calculate the moments from the data points using (6), and reconstruct the distribution using the method we have just formulated.

Figure 1 shows two results of the estimation obtained using polynomial and complex exponential (Fourier) moments. One sees a good agreement between the ME estimate with Fourier moments and the original function, with only slight
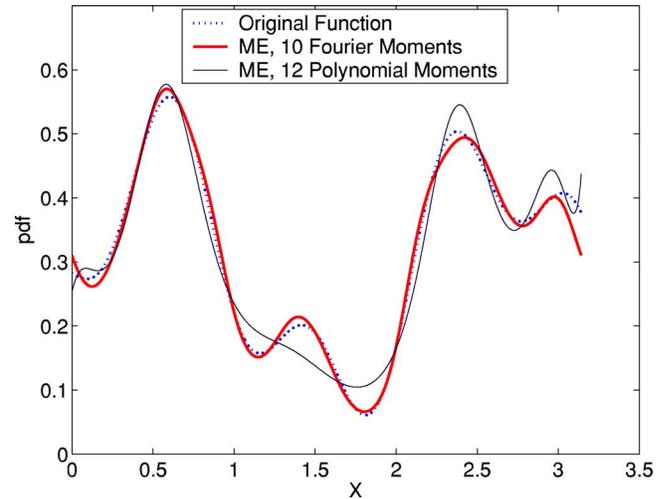


FIG. 1. (Color online) The ME density estimate with polynomial and Fourier moments.

discrepancies around the turning points and near the boundaries. On the other hand, estimation using polynomial moments constraints is unable to resolve several of the sharp gradients. When more polynomial moments are used, the agreement is not improved (see Fig. 2) and even more significantly, the algorithm fail to converge for higher moments. We attribute this to the fact that in solving (13), the polynomial moments lead to a Hilbert matrix known for having high condition numbers (defined as the ratio of largest to lowest eigenvalues). This ill posedness stems from the non-orthogonality of the sequence $\{x^m\}$. Because the different powers of $x$ do not differ greatly from each other at higher moments, the lowest eigenvalues move towards the origin while the larger eigenvalues spread elsewhere [27]. This makes the condition number to increase exponentially with the number of moments, thereby limiting the use of higher moments in practice. Fourier moments do not have such difficulties; their orthogonality leads to well-posed Hermitian-Toeplitz matrix with low condition numbers. We see in Fig. 3 that estimation using higher Fourier moments are stable
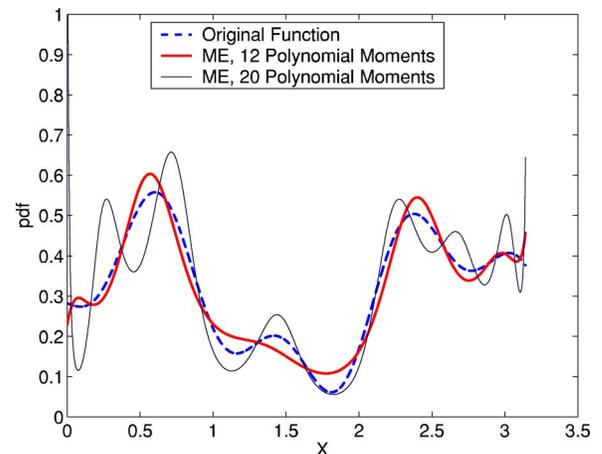


FIG. 2. (Color online) The effect of higher polynomial moments. ME with 12 moments in Figs. 1 and 2 are generated with different random sequences.
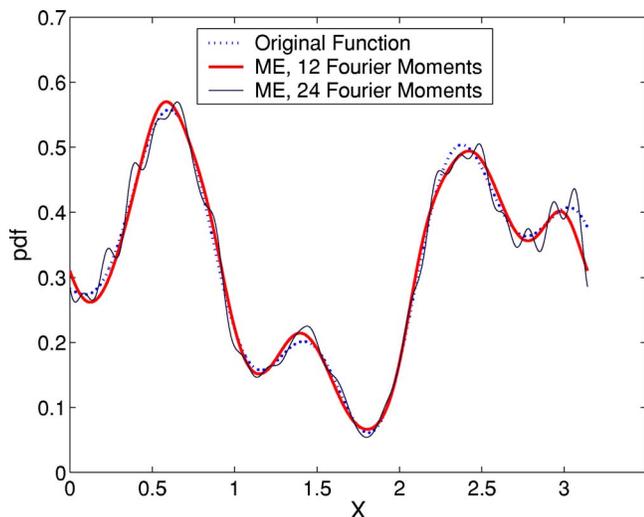
FIG. 3. (Color online) The effect of higher Fourier moments.

though it is susceptible to small high-frequency oscillations. In all the cases studied we find that the Fourier moments exhibit a high degree of numerical stability relative to the polynomial moments. We also point out that a reasonably accurate reconstruction is possible even with a substantially fewer number of particles (of the order of 1000).

Since the kernel method is an alternative approach for density estimation, we show a direct comparison with the ME estimate in Fig. 4. The kernel method which is explained in greater detail in Appendix A uses an optimal window width. We can see that the kernel results which are obtained with 15 000 particles do not perform well near the boundaries. In particular, the boundary values are underestimated, which we attribute to the integral being truncated in the vicinity of the boundary. This density deficiency is a well-known problem associated with the kernel estimates [15].

The conclusions drawn in the one dimensional case studies are generally applicable to higher dimensions as well. With polynomial moments, the ME algorithm is much harder
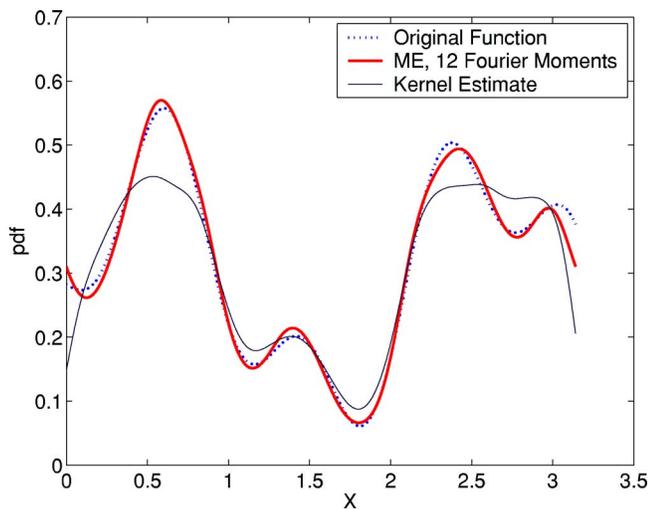


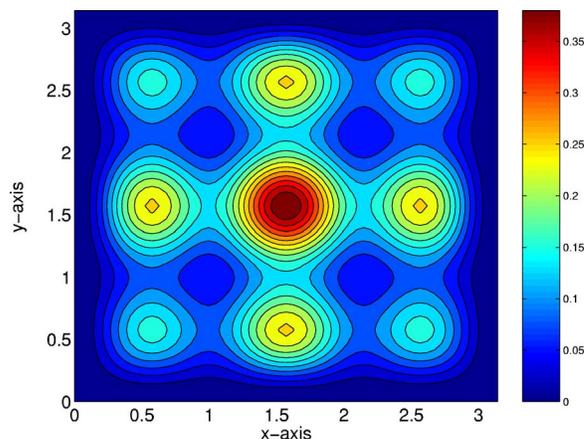FIG. 4. (Color online) A comparison of the ME and kernel density estimates.



FIG. 5. (Color online) The true PDF in two-dimensions.

to converge in two dimensions while Fourier moments do not present this difficulty. We will now illustrate a ME reconstruction in two dimensions. Figure 5 shows the true distribution which is given by

$$\rho(x,y) = \frac{(\sin^2 x + \sin^2 3x)(\sin^2 y + \sin^2 3y)}{\pi^2}. \qquad (17)$$

The normalization constant $\pi^{-2}$ ensures that the integral of the density function is unity. In comparison to the one-dimensional PDF given by (16), the two-dimensional PDF has smaller gradients. Figure 6 shows the ME estimate with 500 particles and four moments in each direction. We see a good agreement with the true distribution and also note that edge effects are minimal in the ME estimate. More realistic fields are estimated from MD simulations and are discussed in the results section. If the gradients are small, then a reasonably accurate reconstruction can be made with relatively few particles (of the order of 1000) and moments ranging from four to six.

### B. Maximum likelihood (ML) velocity and temperature field estimator

We will briefly recall this formulation [13] in preparation for the discussion of results and comparisons. The basic as-
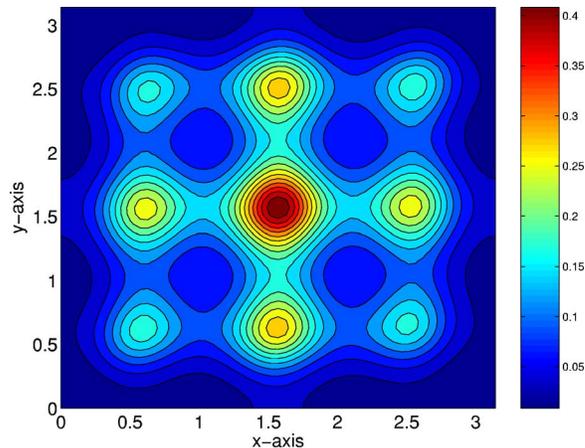


FIG. 6. (Color online) A ME reconstruction in two-dimensions with four Fourier moments.
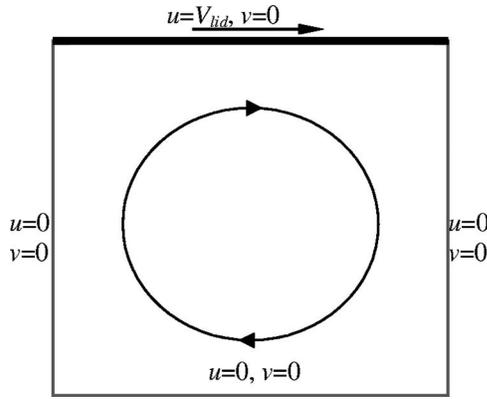
FIG. 7. A schematic of the shear driven flow in an enclosure.

sumption in this parametric approach is that the particles in the simulation follow a local Maxwellian distribution,

$$dP = \frac{1}{[2\pi T(\mathbf{x})]^{D/2}} \exp\left(-\frac{|\mathbf{v} - \bar{\mathbf{v}}(x)|^2}{2T(\mathbf{x})}\right) d\mathbf{x} d\mathbf{v}, \quad (18)$$

where $\bar{\mathbf{v}}$ is the streaming velocity and $D$ is the dimensionality of the system. We choose spatial basis for the field variables $\bar{\mathbf{v}}$ and $T$,

$$\bar{\mathbf{v}}(\mathbf{x}) = \sum_{\mathbf{m}} a_{\mathbf{m}} \varphi_{\mathbf{m}}(\mathbf{x}), \quad (19)$$

$$T(\mathbf{x}) = \sum_{\mathbf{n}} b_{\mathbf{n}} \theta_{\mathbf{n}}(\mathbf{x}), \quad (20)$$

where $\varphi$ and $\theta$ are typically low-order basis functions. The likelihood, $P$, for the particles to have certain positions and velocities is given by:

$$P(\{a_{\mathbf{m}}, b_{\mathbf{n}}\}) = \prod_{i=1}^{N} dP(\mathbf{X}_i, \mathbf{V}_i). \quad (21)$$

The most probable field distribution for velocity and temperature in the parameter space $\{a_m, b_n\}$ is computed by maximizing the logarithm of this likelihood function [13].

### III. THE MODEL PROBLEM

We investigate the effectiveness of the statistical field estimators by considering the behavior of a two-dimensional shear or lid-driven flow in an enclosure as determined by molecular dynamics simulation and by the continuum Navier-Stokes equations. The idea is to benchmark the distributions obtained from the field estimators against the continuum results.

In the present problem shown schematically in Fig. 7, the top boundary moves at a constant horizontal velocity $V_{\text{lid}}$. For the continuum description, the velocity components at all other wall boundaries are set to zero (no-slip condition). For the MD simulation, the boundaries consist of a few layers of solid atoms; the top array, acting as a lid, is allowed to slide at a constant velocity.

Driven flow in an enclosure exhibits a variety of complex hydrodynamic behavior such as eddies, secondary flows, in-

stabilities, and bifurcations [28]. The incompressible continuum description does not have a closed form solution though for creeping flows, a series solution to a biharmonic stream-function equation is possible. An infinite sequence of secondary vortices of diminishing size exists at the lower corners even for arbitrarily small Reynolds (Re) numbers [28]. At higher Re the secondary vortices grow in size and strength; detailed prediction of flow patterns is possible only through numerical simulations.

Because of its rich features, driven flow in a cavity is a standard benchmark problem for testing continuum numerical schemes. It is also a problem where the boundaries are well-defined without ambiguity [28]. Unlike in Coutte and Poiseuille flows where relatively intricate inflow and outflow boundary conditions are required for MD, flow in a driven enclosure poses no such difficulties. This type of flow therefore, is well suited for comparing MD and continuum results.

### A. MD simulation

MD simulations are performed with a two-body short-smooth potential [29]. The thermodynamic states generated with this model are discussed in Appendix B. The shear viscosity is calculated using the Green-Kubo linear response theory as the time integral of the stress autocorrelation function which is obtained from MD simulations performed on a square domain with periodic boundary conditions. The expression for shear (dynamic) viscosity is [30]

$$\mu = \frac{V}{k_B T} \int_0^\infty dt \langle \sigma_{xy}(t) \sigma_{xy}(0) \rangle, \quad (22)$$

where $\sigma_{xy}$ represents the shear stress tensor, $T$ is the temperature, and $V$ is the volume (area) of the simulation cell. The angular brackets denote a time average over a sufficiently large number of independent samples. For a two-body potential, the shear stress tensor, just like in continuum Newtonian flow, is symmetric.

Equilibrium simulations for determining the viscosity are conducted with approximately 3000 atoms in a square domain measuring 70 by 70 (reduced) units, with runs typically of 9 million time steps with a step size of 0.001. The integral in (22) is evaluated using fast Fourier transform as explained in [31]. For the same density and temperature, the calculated viscosity is seen to be insensitive to the system size.

To perform the nonequilibrium simulations with shear flow, fluid atoms are confined in a square enclosure bounded by a few layers of solid atoms on each side which constitutes the physical boundary of the simulation cell. A liquid or gaseous state is created by randomly removing a fraction of the fluid atoms from the computational domain and allowing the system to equilibrate. The MD simulation cell viewed with *AtomEye* [32] is shown in Fig. 8.

Shearing of the system is simulated by dragging the top layer at a constant velocity. Due to the external work done on the system the internal energy of the fluid atoms increases with time. A first order thermostat is used to keep the system at constant bulk temperature. All our simulations are performed under the conditions where the local flow velocities are smaller than the thermal velocities, so the system tem-
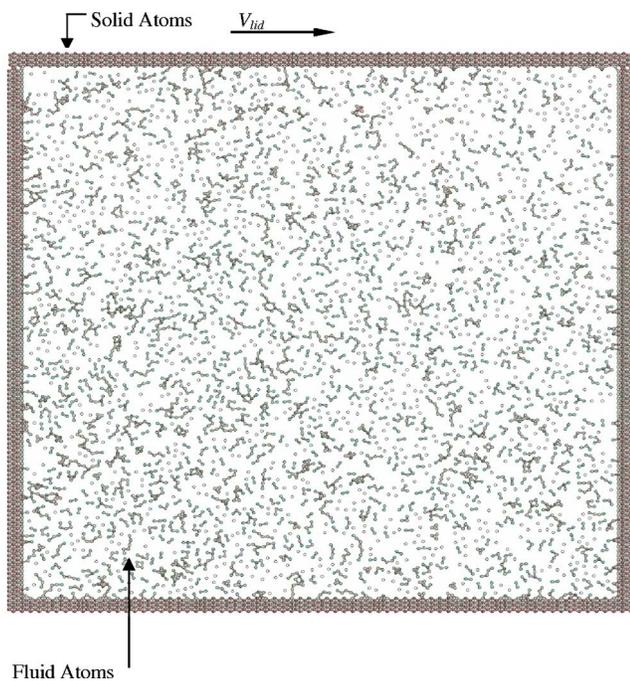
FIG. 8. (Color online) The MD simulation cell. Atoms and bond connectivities are shown.

perature is dominated by the latter. The use of a thermostat is appropriate because the hydrodynamic behavior, which is of primary interest in this study, is not sensitive to the method of heat removal even for reasonably large shear rates [18,33]. Additionally, we find that the long wavelength modes are largely independent of the details of the interatomic potential used in the simulation, as also noted previously [34].

To integrate the equations of motion we use Gear's fifth order predictor-corrector scheme. $O(N)$ efficiency is achieved by using a cell list method, the details of which are given in [13]. The interactions at the boundary are treated through direct fluid wall interactions. Simulations are carried out for two cells, measuring $200 \times 200$ and $100 \times 100$ (reduced) units respectively. At $(\rho, T) = (0.48, 1)$, the larger system consists of 19 257 fluid atoms and four layers of solid atoms with a total of 23 105 atoms. For the same thermodynamic state, the smaller system comprises of 4814 fluid atoms and 1952 wall atoms. Due to the slow convergence of the long wavelength fluctuations, simulation runs are typically executed for 20 to 30 million time steps. Further details are given when the results are discussed in Sec. IV. We also note that Koplik and Banavar [18] have reported a similar MD study focused on the slip behavior near the moving solid boundary.

### B. Continuum simulation

The fields generated by the estimators are expected to reflect the slow and long wavelength modes of the atomic motion in the MD simulation [35]. A direct way to check the fidelity of the estimation is to compare the fields obtained with the results given by the Navier-Stokes (NS) equations [36]

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \tag{23}$$

$$\frac{\partial}{\partial t}(\rho \mathbf{v}) + \nabla \cdot (\rho \mathbf{v} \mathbf{v}) = \nabla \cdot \tau + \rho \mathbf{f} \tag{24}$$

$$\frac{\partial}{\partial t}(\rho e) + \nabla \cdot (\rho e \mathbf{v}) = \nabla \cdot (\tau \cdot \mathbf{v} - \mathbf{q}) + \rho \mathbf{f} \cdot \mathbf{v}, \tag{25}$$

where $\tau$ is the stress tensor, $\mathbf{f}$ is the external force, and $\mathbf{q}$ is the heat flux vector. The above equations are closed by invoking Newtonian relations concerning stress and strain rate, Fourier law of heat conduction, Stokes hypothesis connecting the viscosity coefficients, and the appropriate equations of state.

Since the NS equations for the driven cavity do not have a closed form solution we resort to numerical simulations. A number of solvers have been developed in the past to simulate the flow features of the driven cavity, almost all based on the assumption of incompressible flow. We employ a numerical algorithm involving vorticity and stream function as described by Pozrikidis [37] to simulate the incompressible flow fields. Further details are given in Appendix C.

MD simulations, in general, correspond to neither incompressible flow nor Newtonian behavior, especially at shear speeds comparable to the thermal velocity. To evaluate the effects of compressibility on the flow field, we have simulated the flow of a compressible fluid using Fluent® [38] which discretizes Eqs. (23)–(25) into finite volumes and solves the resultant equations with an algebraic multigrid solver. To our knowledge, there are no published results on compressible flow in a driven confined cavity.

To link the MD and continuum simulations one considers two measures, the Reynolds number (Re), the ratio of inertial to viscous force,

$$\mathrm{Re} = \frac{\rho V_{\mathrm{lid}} L}{\mu}, \tag{26}$$

where $\rho$ is the mass density, $\mu$ is the dynamic viscosity, and $L$ is the linear dimension of the square cavity, and the Mach number based on lid velocity,

$$M = \frac{V_{\mathrm{lid}}}{c}, \tag{27}$$

where $c$ is the sound speed sound measured with the properties along the lid. For incompressible continuum simulations, only the Reynolds number appears in the problem. For the MD simulations, incompressibility is verified by evaluating the spatial variations of the density and Mach number fields.

The continuum field equations also tacitly assume that the ratio of mean free path of the molecules to the characteristic length (Knudsen number) is small ($\ll 1$) and the shear velocities are much smaller than the thermal velocities of the atoms so that local thermal equilibrium is always maintained at all times. In our MD simulations, the fluid density and system size are such that the Knudsen number is less than 0.05, which constitutes a reasonable condition for the continuum approximation. The shear rates however, are of the order of
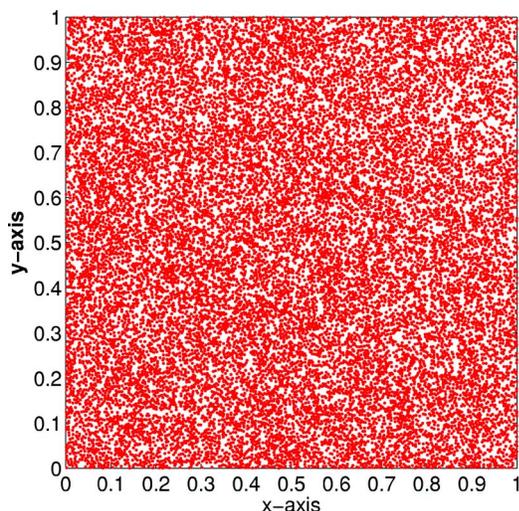
FIG. 9. (Color online) The scatter plot of fluid atoms, 200 × 200 domain, lid velocity=0.1.



FIG. 10. (Color online) The ME density field, single snapshot, 4 moments with a 200 × 200 domain, lid velocity=0.1.

0.1, much higher than what is realizable in the macroscopic conditions. So in addition to Reynolds and Mach numbers, we also consider the ratio of lid velocity to the mean thermal velocity ($V_{lid}/V_{th}$) in interpreting the comparisons.

## IV. RESULTS

### A. Density field estimation

We compare the density fields obtained from the particle data of two MD simulations at the same thermodynamic state. For the larger simulation cell with a lid velocity of 0.1 (reduced units), the Reynolds number is 30, the Mach number is 0.077 and the ratio of lid velocity to thermal speed ($V_{lid}/V_{th}$) is 0.07. For the smaller system with a higher lid speed of 0.68, the corresponding numbers are 100, 0.523, and 0.48 respectively. We expect that in the first case the evolution of the mean field will be largely independent of the molecular motions, whereas in the second case there can be appreciable dynamical coupling.

Figure 9 shows the scatter plot of the atoms for the larger system (19 257 fluid atoms). The particles are more or less evenly spread across the domain although upon closer scrutiny one can discern a slight under-population in the upper right region and a slight over-population in the center and lower region. We anticipate a near uniform density because the thermal speed (1.4) is more than an order higher than the lid velocity (0.1) and the Mach number is small.

Figure 10 depicts the density field distribution estimated using the maximum entropy method with four Fourier moments while the average based on 169 bins is shown in Fig. 11. The field representation with ME captures the essential characteristics of the underlying distribution. While the bulk of the domain is clearly close to uniform (within 15%), the nonuniform regions are also quite evident. The essential characteristics are consistent with the scatter plot in revealing a lower density at the upper right corner and a higher density near the central and lower regions. The number of moments controls the smoothing effect in the present method; using
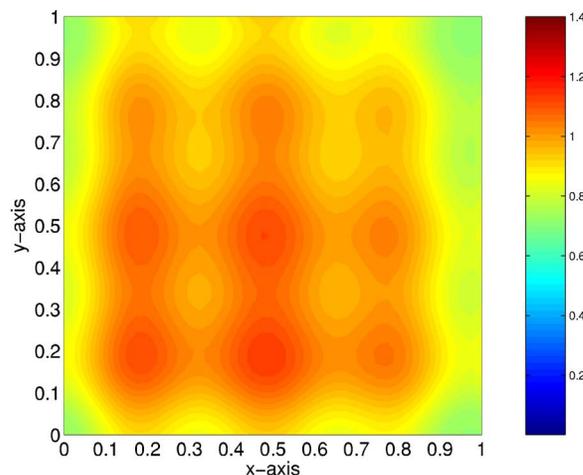
more moments tends to resolve the sharper gradients better.

There is no formal way to determine the optimal number of moments *a priori*. Since all nonparametric estimations involve at least one or two parameters, the best estimate is often a matter of intuition based on what one knows about the distribution being estimated [21]. In our study, the scatter plots, being a reflection of the molecular data with no analysis, serve as a guide in assessing the fidelity of the estimator. The bin average, as expected, shows a very jagged topology with large and uneven gradients that can lead to numerical artifacts in multiscale simulations. Similar lack of smoothness can also be noted in [34]. Note that we have used a single snapshot to compare the fields. It is fair to say that a reasonable bin average needs an average over several hundred such snapshots to smooth out the sharper gradients. We have not observed a dramatic improvement in the smoothness for such a time averaged bin field. This aspect is brought in more detail in the discussion for the next density field where the gradients are more discernible.

Next we show the kernel estimate in Fig. 12. We observe a markedly lower density region along the borders and near
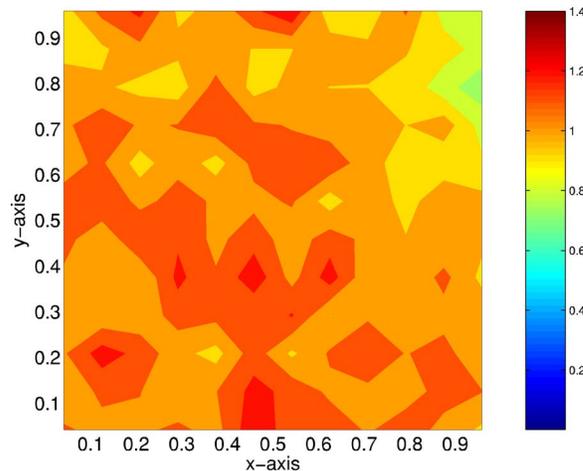


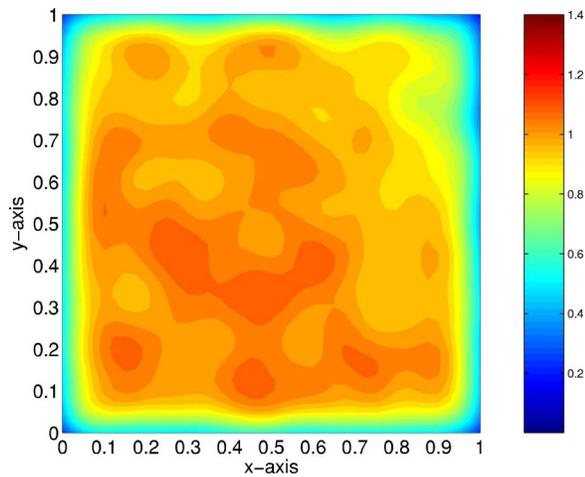FIG. 11. (Color online) The bin average, single snapshot with a 200 × 200 domain, lid velocity=0.1.

FIG. 12. (Color online) The Gaussian kernel density field, single snapshot, $200 \times 200$ domain, lid velocity=0.1.



FIG. 14. (Color online) The Gaussian kernel density field, single snapshot, $100 \times 100$ domain, lid velocity=0.68.

the leading edge (top right corner) which is not seen in the scatter plot. This deficiency occurs because the kernel for the particles gets truncated at the system boundary which results in a one-sided contribution to the total density estimate. On the other hand, the ME distribution does not suffer from this edge artifact. The edge deficiency is known to affect the accuracy of SPH simulations which has led to the introduction of virtual particles at the boundaries for remediation [15].

We now consider the smaller system with a higher lid speed of 0.68 corresponding to a Reynolds number of 100 and a Mach number of 0.53. The ratio of lid velocity to the thermal velocity is 0.48. The density fields estimated by ME (with 5 moments) and kernel averaging are shown in Figs. 13 and 14, respectively. The corresponding scatter plot is given in Fig. 15.

The lid velocity, now a significant fraction of the thermal velocity, produces a region of high stress near the leading edge (upper right corner). The flow in the cavity is characterized predominantly by a single large vortex moving in the clockwise direction (discussed later). Due to the asymmetry
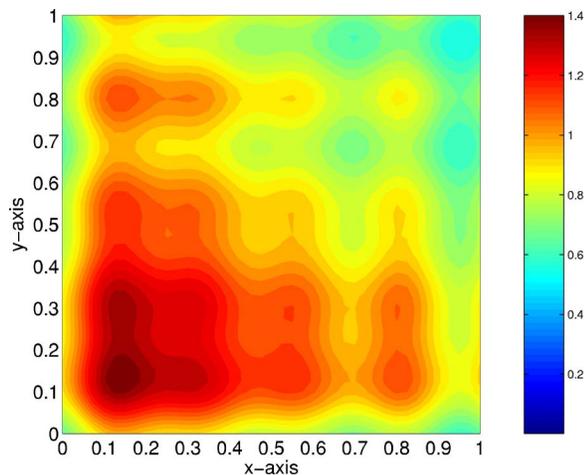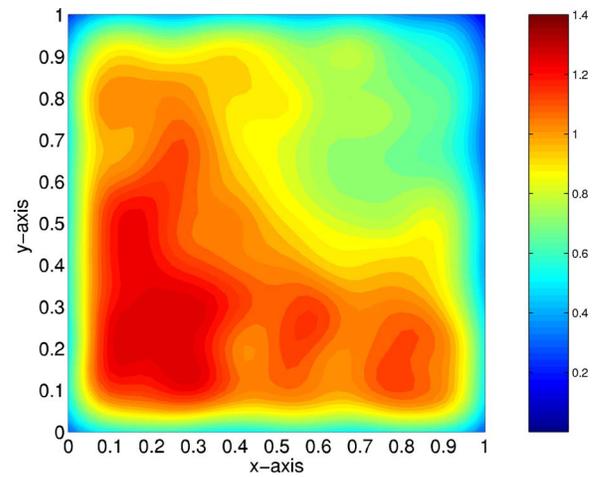
of the stress field in the vicinity of the leading corner, the particles migrate towards the regions of lower stress, thereby causing a higher density at the lower left region. This aggregation represents shear localization that is induced by the confining walls.

From a series of simulations at different densities and temperatures, we find that this effect of density pileup is roughly proportional to $V_{\mathrm{lid}}/V_{\mathrm{th}}$ and not correlated with the Reynolds number. This also indicates a breakdown of dynamic similarity (where the nondimensional solutions are identical) when the ratio, $V_{\mathrm{lid}}/V_{\mathrm{th}}$ is $O(1)$. Even with this discernible density gradient, as we will discuss later, the velocity distributions appear to be remarkably similar to those from the continuum simulations. Additional results including structural and dynamical correlation functions under different shear rates will be reported in a future publication.

Both ME and kernel methods give the highest density in the lower left region and lowest density in the upper right corner. As also seen in the larger system, the kernel estimated field falls off near the boundaries, while the ME field does



FIG. 13. (Color online) The ME density field, single snapshot, 5 moments, $100 \times 100$ domain, lid velocity=0.68.
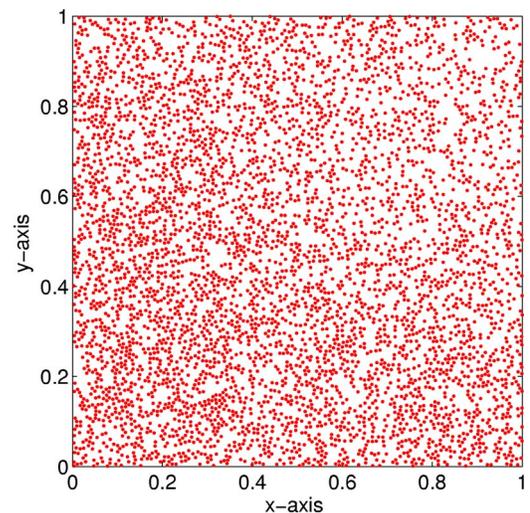


FIG. 15. (Color online) The scatter plot of fluid atoms, $100 \times 100$ domain, lid velocity=0.68.
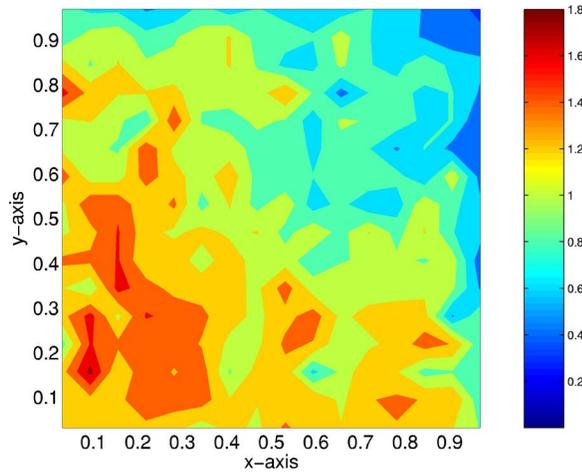
FIG. 16. (Color online) The bin average, single snapshot, 100 ×100 domain, lid velocity=0.68.



FIG. 18. (Color online) The time-averaged density profile at $x$ =0.5.

not have this difficulty. In the interior, the two estimated fields appear to be quite similar.

Figure 16 shows a bin average of the same simulation data with 289 bins based on a single snapshot. In Fig. 17, we depict the bin density field averaged over 1000 snapshots over a period of 20 000 time steps. The density field over this period does not change appreciably and the time averaged ME and kernel fields are similar to those depicted in Figs. 13 and 14.

While we see a noticeable improvement in the time aver-aged bin field, the jaggedness in the topology still persists and the sharp gradients do not altogether disappear simply by extending the duration of time averaging. Theoretically, a smooth bin average is only possible when there are a large number of particles in addition to a sufficiently large number of time-averages. From a multiscale simulation point of view, this gives a significant advantage to the ME method because for a similar degree of smoothness, bin averaging requires a larger number of particles in addition to a larger number of time averages. Alternatively, interpolation can be used to smooth out the jaggedness in the bin field but will
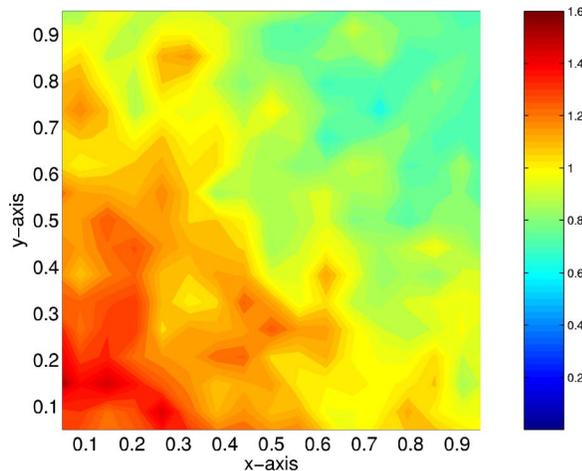
require additional and nontrivial effort. The ME method is naturally extendable to higher dimensions and as shown in our studies, it generates smooth and reasonably accurate fields with a limited number of particles and time averages. The time averages for the ME method are dictated only by the period that is necessary to arrive at the correct physics. Generally, this period amounts to a few thousand iterations in a molecular dynamics simulations where the higher moments of the distribution "heals" and becomes functionals of the lower moments that correspond to the macroscopic field variables [8]. No additional averages are required for smoothing purposes in the ME method. We show the effect of time averages on the density profile at $x$=0.5 with the three methods in Fig. 18. The bin and kernel fields are aver-aged over 1000 sets over a period of 20 000 times. The ME estimate which is averaged over just 100 sets in the same time period is smooth and shows the same trend as that is seen in the kernel and the bin averages.

It is worthwhile to note that local features are not always well-defined by the ME estimator. The generated distribution displays a somewhat exaggerated undulation, though of small magnitude, which is not evident in the scatter plot. These oscillations are due to the dominance of the lower moments of the distribution. Theoretically, if the first two polynomial moments are employed, then the resulting ME distribution is an exponential function [24]. If three moments are prescribed, then the ME distribution will be a Gaussian. To first-order, this is true for Fourier moments too. When a small set of moments are employed, such as four or five (as in our simulations), the resulting distribution retains some of the Gaussian characteristics. As the number of moments in-creases, the Gaussian smoothing becomes less prominent. It is also important to note that these undulations are small in magnitude and smooth, as evident in Fig. 18. Since ME is able to capture the key features of the distribution quite well, the minor waviness will not be a hindrance in multiscale simulations where both accuracy and smoothness are prereq-uisites.

The present formulation of ME method uses a set of in-tegrated moment constraints and does not employ any local



FIG. 17. (Color online) The bin field averaged over 1000 snap-shots, 100×100 domain, lid velocity=0.68.

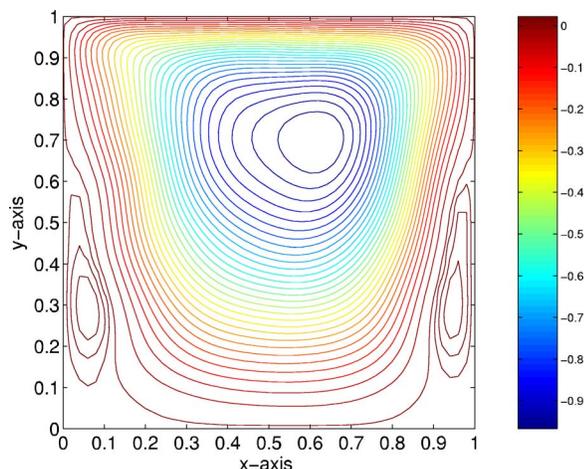FIG. 19. (Color online) The generalized streamlines with MD simulation, lid velocity=0.68, Re=100.



FIG. 20. (Color online) The streamlines from the incompressible Navier-Stokes equation, Re=100.

particle information *per se*. In Sec. VII, we discuss a possible improvement to the present formulation to capture the local information. The numerical cost of the current ME method is $O(N^2)$ but with a fast Fourier transform (FFT) implementation, the cost can be brought down to $O(N \log N)$. The convergence of Newton-Raphson scheme is fairly rapid, typically takes less than 50 iterations for the relative error in entropy to decrease by six orders.

### B. Velocity and streamline fields

We demonstrate the effectiveness of ML field estimators by comparing the streamlines and the velocity fields with the continuum results. The conventional approach for calculating the streamlines from the velocity and density fields involves a direct integration of the mass flux [30]

$$\Omega(x,y) = \int \rho(x,y)(v_y dl_x + v_x dl_y). \qquad (28)$$

The integral is typically performed by assigning an arbitrary value to the stream function at a grid point and then performing the line integral to evaluate the value at a different grid point. We have used an alternate approach to increase the smoothness by first transforming the mass flux into a generalized vorticity field and then solving the vorticity-streamfunction Poisson equation. This method gives very smooth streamlines and is a definite improvement over (28). The details of the formulation are given in Appendix D. The continuum results are obtained directly from the incompressible streamfunction as described in Appendix C.

We compare the generalized streamlines obtained from the MD results with those from the continuum calculations, in Figs. 19 and 20, respectively. The difference between two adjacent streamfunctions, by definition, is a measure of the mass flow rate between them. No fluid can cross the streamlines because the velocity vector is tangent to the surface. In a confined enclosure, all streamlines therefore should close onto themselves to maintain mass conservation. We see that the MD streamlines show this behavior very clearly, which
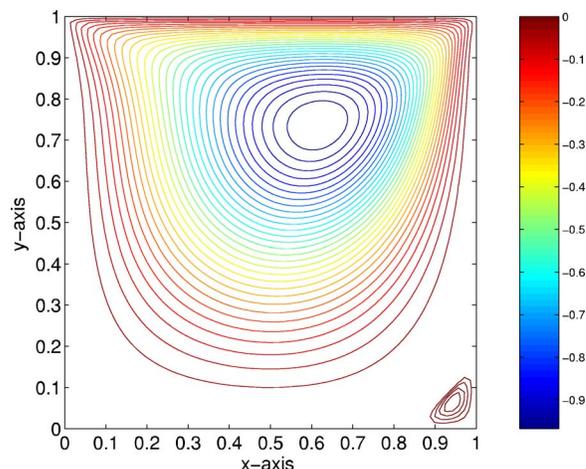
may be taken as a test of the efficacy of our vorticity-streamfunction formulation. Direct integration which is cumbersome may lead to artifacts as observed in [18,19,39]. Overall, the streamlines generated by the MD data, in spite of having a discernible density gradient, match well with those obtained from the continuum description. This is reasonable when the local flow velocities are smaller than the thermal velocities and the simulation conditions clearly indicate that of the hydrodynamic limit. In both results, the eye of the vortex is shifted towards the right, in the direction of the lid velocity, at about the same horizontal position. From the extent of the streamlines we see the flow region covers most of the computational domain in the MD simulation, while in the continuum calculation, very weak anticlockwise eddies appear near the lower corners. In the MD simulation the strength (numerical value of the streamfunction) of the secondary vortices is much larger because of explicit fluid-wall interactions.

Figure 21 depicts the horizontal velocity field given by the ML estimator. The fields generated with the continuum simulation and the kernel estimator are given in Figs. 22 and
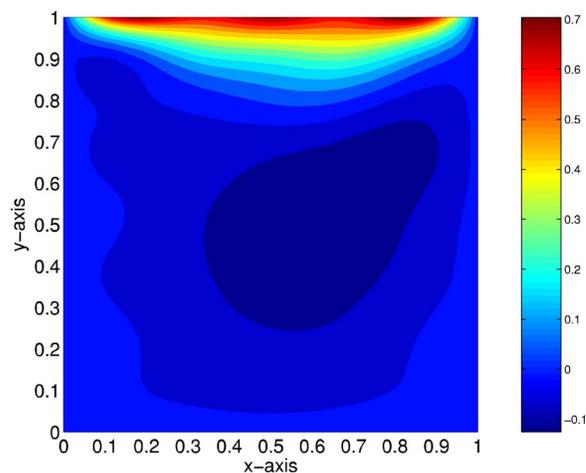


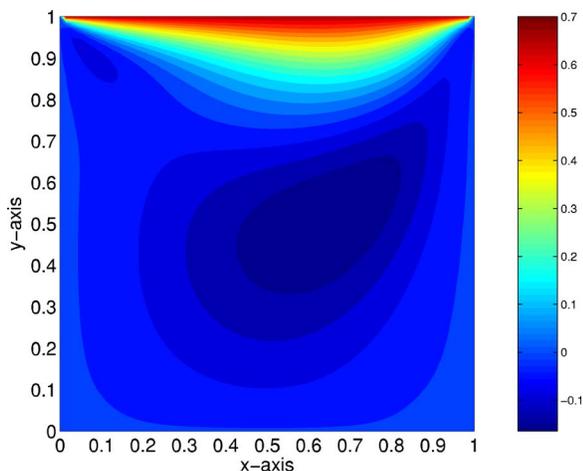FIG. 21. (Color online) The horizontal velocity field with ML estimator, lid velocity =0.68, Re=100.

FIG. 22. (Color online) The horizontal velocity field with the NS simulation, lid velocity =0.68, Re=100.



FIG. 24. (Color online) A comparison of the horizontal velocity at the vertical midplane, Re=100.

23, respectively. The ML field is remarkably similar to that of the continuum. Adjacent to the lid, the former shows a flow velocity nearly the same as that of the lid, except near the corners where it tapers off to zero. This indicates that the commonly invoked no-slip boundary condition is valid for the central regions but not near the corners, an observation which also has been noted by Koplik [18]. At higher lid speeds, substantial slip is seen between the moving boundary and the nearby fluid (results not presented). With the kernel average, we note that while the general flow features are reasonable, the velocity gradients near the lid are more diffuse. In addition we see a velocity slip behavior near the upper corners which is not observed in the ML estimate. We believe this to be an artifact.

In Fig. 24 we compare the horizontal component of the velocity along the vertical midplane obtained from the MD data and continuum calculations. In Fig. 25 we give the vertical velocity at the horizontal midplane. Note that the ML velocity profiles are derived through the estimator without any additional smoothening. The ML estimations are in overall agreement with the continuum description. The kernel
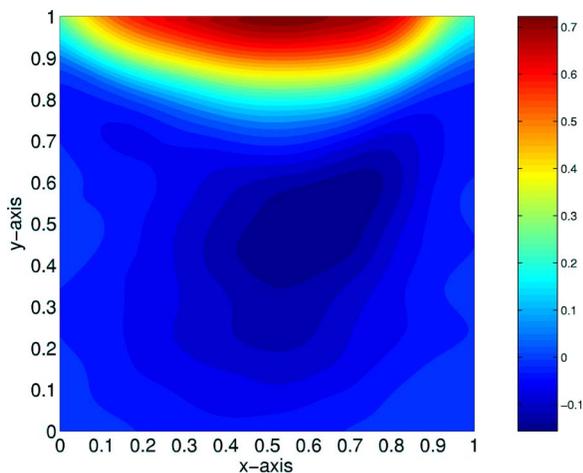
estimates compare favorably with the ML estimate in the interiors. As before, the kernel estimation suffers from numerical artifacts at the boundaries, the effect which is clearly seen in Fig. 25 in the nonzero values at the right and left boundaries. Velocity slip at the boundaries in this case is clearly numerical in origin; it should not be taken as an indication of slip phenomena at the molecular scales. ML estimate does not suffer from such edge artifacts which we believe is a key strength of the statistical field estimators.

Figure 26 shows the vertical velocity profile at the horizontal midsection for ML and bin estimate after 150 000 MD time steps. Evidently, ML takes a fewer number of iterations to produce a reasonably accurate solution that is also smooth. We also add that in multiscale simulations such as the "equation-free" method, the MD simulators are envisioned to run far fewer time steps than 150 000. The idea is to arrive at the steady state solution using the gradient information in the fields at regular intervals.

A lid speed of 0.68 is a significant fraction of the acoustic speed, taken as $\sqrt{\gamma T}$ by assuming that the fluid behaves like



FIG. 23. (Color online) The horizontal velocity field with the kernel estimator, lid velocity velocity=0.68, Re=100.
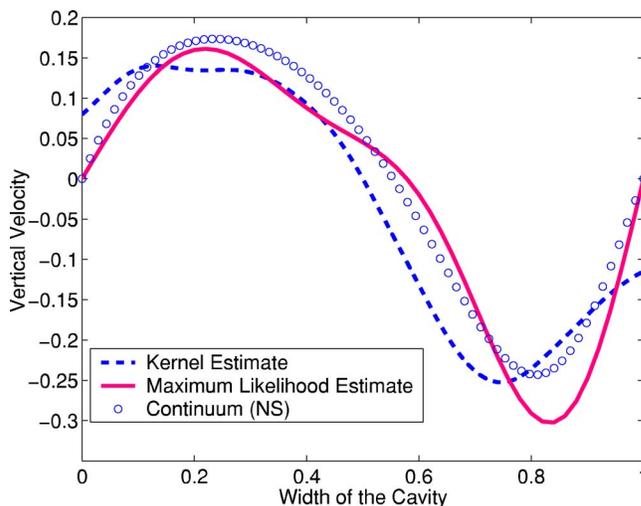


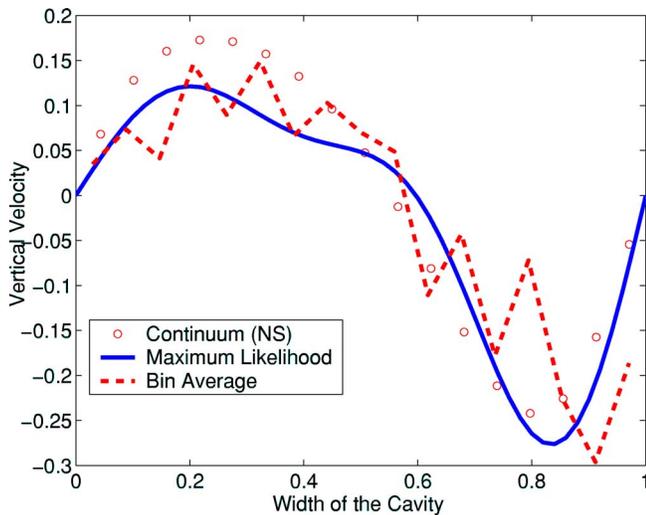FIG. 25. (Color online) A comparison of the vertical velocity at the horizontal midplane, Re=100.

FIG. 26. (Color online) A comparison of time averaged ME and bin field estimates. Lid velocity=0.68, Re=100, $N_T$=150 000.



FIG. 28. (Color online) The local mach numbers with the MD simulation. Lid velocity=0.68, Re=100.

a monatomic ideal gas. At a temperature of 1.0, the acoustic speed is approximately 1.3. Therefore, the highest Mach number in the system is roughly 0.53. Generally, if the Mach number is greater than 0.2, the flow is considered to be compressible. To evaluate whether the compressibility has any appreciable effect on the continuum field variables, we study the flow of a compressible fluid in an enclosure using the CFD software Fluent® [38]. Simulations are performed with a fine 150 by 150 grid with all the boundary walls kept at a constant temperature of 300 K. The reference pressure is set at the moving lid and the properties of air are assumed for the medium. Using an implicit coupled algorithm, the mass, momentum, and energy equations [Eqs. (23)–(25)] are solved simultaneously and accelerated with algebraic multigrids.
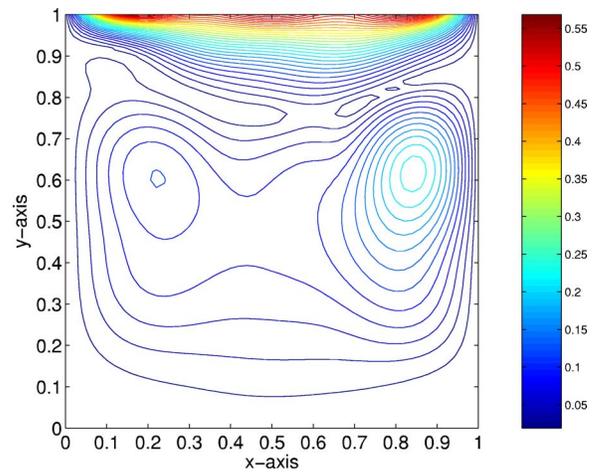
The Mach number shown in Fig. 27, indicates that that the bulk of the flow remains incompressible, with the compressible regime limited to a thin region along the lid. We also note that the velocity fields are only marginally affected by the effect of compressibility (results not shown). The local Mach number profiles with MD as seen in Fig. 28 are somewhat similar to those from the continuum with the high Mach number regions confined near the lid. The relative closeness of the MD Mach number and the velocity contours with the continuum results indicates that dynamic similarity is partially satisfied at the molecular scales for the shear rates considered in our study.

## V. SUMMARY AND DISCUSSION

Bin averaging is the work horse for most MD simulations where the continuum fields are to be constructed from the
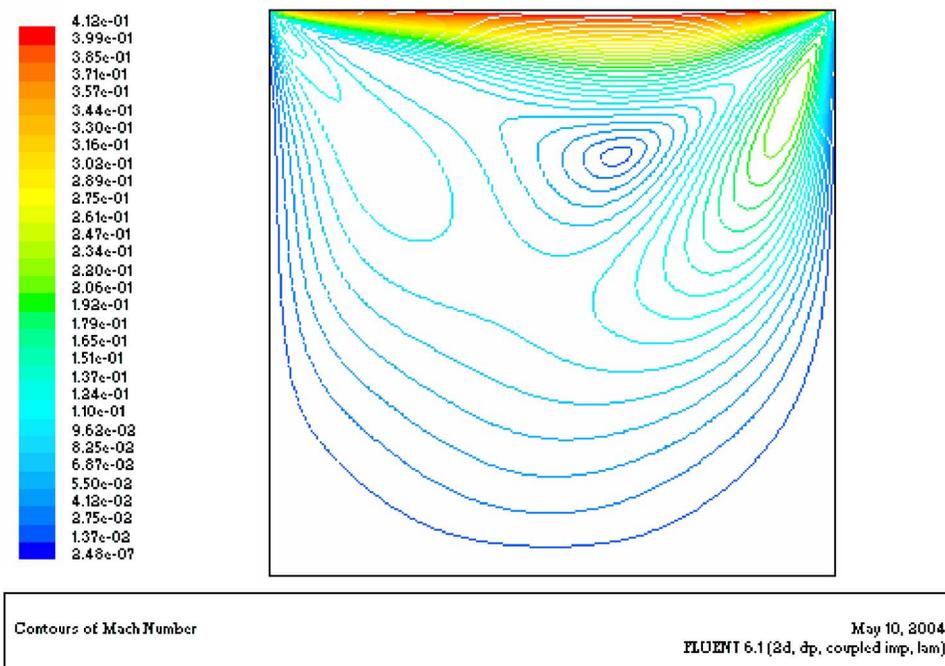


FIG. 27. (Color online) The local Mach numbers with the compressible NS simulation (Fluent Inc.)®, Re=100.

discrete particle data. Poor resolution and statistical noise make the bin method an unsatisfactory choice for performing complex multiscale simulations where smoothness, in addition to accuracy, is a necessity over a relatively short period of time. To avoid the limitations of bin averaging, we have presented a set of statistical field estimators which generate smooth and accurate fields. For spatial density distributions, we have proposed a nonparametric estimator based on the maximum-entropy (ME) principle. The ME estimator identifies the least biased distribution that is consistent with the prescribed set of constraints which are in form of moments of the distribution. We show that we can calculate the moments from the particle data without further assumptions. The results from an MD simulation of a shear driven flow in an enclosure, show that the ME estimation is superior to the bin average. In comparing the ME estimator with the kernel average, we observe that the kernel method systematically underestimates the density near the boundaries while ME method does not exhibit this artifact.

We have generated the streamline and velocity fields from maximum likelihood technique (ML) which again are appreciably smoother than the bin average. We have also compared the ML velocity field with the kernel average. While the interior flow field is satisfactorily resolved, the edge artifacts near the boundaries pose a difficulty for the kernel average. The velocity slip at the boundaries in the kernel method is clearly numerical in origin and it should not be construed as an indication of slip phenomena at the molecular scales. We have also discussed a method for calculating the streamlines by solving the generalized vorticity-streamfunction Poisson equation. This approach not only gives closed streamlines but also adds an extra derivative which imparts additional smoothness.

We have uncovered an interesting form of shear localization which is induced by the confining boundaries of the system. From a series of simulations at different densities and temperatures, we find that this localization is roughly proportional to $V_{\text{lid}}/V_{\text{th}}$ and not correlated with the Reynolds number. Even with a discernible density gradient in the flow field, the velocity distributions appear to be remarkably similar to those from the continuum simulations. Dynamic similarity, however, breaks down completely when the ratio of $V_{\text{lid}}$ to $V_{\text{th}}$ is $O(1)$.

The proposed ME density estimator employs a set of integrated moment constraints. This may be classified as a "global" approach in contrast to the kernel method which makes direct use of local information. For field representation where there are few data points, it is advantageous to include local particle information directly in the ME estimator. One way to do this is to break down the domain into regions such that each region contains one particle. Then the ME principle can be applied to each region and a piecewise-continuous estimate can be made [24]. Additional smoothness can be imparted by employing higher order moments in the local spacing. This approach is feasible for one dimension but in two and three dimensions proper spacing is not easily derivable because of lack of order statistics in higher dimensions. Voronoi and Delaunay tessellations [40] provide a reasonable procedure for characterizing the local spacing around a particle. Then the density distribution can be constructed by applying the principle of maximum entropy to each tessellation. This way ME method can access the local information while at the same time provide adequate smoothness.

## APPENDIX A: KERNEL FIELD ESTIMATORS

Kernel averaging is a widely used technique in nonparametric density estimation [20,21]. In this method, the distribution $\rho(\mathbf{x})$ is the average of kernel or smoothing functions centered on the given particle positions $\mathbf{X}$. The particle density distribution $\rho(\mathbf{x})$ of $N$ particles as given by (5) is

$$\rho(\mathbf{x}) = \frac{1}{N}\sum_{n=1}^{N} \delta(\mathbf{x} - \mathbf{X}_n). \qquad (A.1)$$

In the kernel method we substitute a smooth function for the delta function to obtain

$$\rho(\mathbf{x}) = \frac{1}{N}\sum_{n=1}^{N} K(\mathbf{x} - \mathbf{X}_n, h_n), \qquad (A.2)$$

where $K(\cdot)$ is the smoothing kernel function and $h$ is the smoothing scale (window width) which determines the support of $K(\cdot)$. Note that the smoothing scale can be different for each particle. $K(\cdot)$ is usually taken to be a symmetric function even though this is not a strict requirement.

To represent a *bona fide* density distribution, the kernel should have the following properties [15]:

$$\int_{\Omega} K(\mathbf{x} - \mathbf{X}, h)d\mathbf{x} = 1, \qquad (A.3)$$

$$K(\mathbf{x} - \mathbf{X}, h) \geq 0, \qquad (A.4)$$

$$\lim_{h \to 0} K(\mathbf{x} - \mathbf{X}, h) = \delta(\mathbf{x} - \mathbf{X}). \qquad (A.5)$$

Smoothness is a desired characteristic of the kernel function. In this study, we utilize a Gaussian kernel with zero covariance which is given by,

$$K(\mathbf{x}) = \frac{1}{(2\pi)^{D/2}} \exp\left(-\frac{1}{2}\mathbf{x}^{\mathbf{T}}\mathbf{x}\right). \qquad (A.6)$$

The density distribution with a single smoothing length is evaluated as

$$\rho(\mathbf{x}) = \frac{1}{Nh^D} \sum_{n=1}^{N} K\left(\frac{\mathbf{x} - \mathbf{X}_n}{h}\right). \quad (A.7)$$

Other kernel functions such as Epanetchnikov and Biweight [21] are available but offer no significant computational advantage. Further, very small support localizes the field which is undesirable in the regions of low density. The optimal smoothing scale is calculated by minimizing the asymptotic integrated square bias and integrated variance [41]. For most cases studied in this paper, this window width has been found to be optimal in the sense that it resolved the sharp gradients without causing large undulations.

### Kernel velocity and temperature field estimator

The kernel estimation method can be extended to evaluate the momentum and temperature fields. Substituting a smooth function for the delta function in (7) we get

$$\mathbf{p}(\mathbf{x}) = \frac{1}{N} \sum_{n=1}^{N} \mathbf{V}(\mathbf{X}_n) K(\mathbf{x} - \mathbf{X}_n, h_n). \quad (A.8)$$

Since the velocity field is given by

$$\mathbf{v}(\mathbf{x}) = \mathbf{p}(\mathbf{x})/\rho(\mathbf{x}), \quad (A.9)$$

the kernel estimate for the velocity takes the form

$$\mathbf{v}(\mathbf{x}) = \frac{\sum_{n=1}^{N} \mathbf{V}(\mathbf{X}_n) K(\mathbf{x} - \mathbf{X}_n, h_n)}{\sum_{n=1}^{N} K(\mathbf{x} - \mathbf{X}_n, h_n)}. \quad (A.10)$$

The local temperature field can now be calculated as

$$T(\mathbf{x}) = \frac{\frac{1}{2} \sum_{n=1}^{N} |\mathbf{V}(\mathbf{X}_n) - \mathbf{v}(\mathbf{x})|^2 K(\mathbf{x} - \mathbf{X}_n, h_n)}{\frac{D}{2} \sum_{n=1}^{N} K(\mathbf{x} - \mathbf{X}_n, h_n)}. \quad (A.11)$$

### APPENDIX B: SHORT-SMOOTH (SS) POTENTIAL

The atoms interact through a short-smooth (SS) potential [29] which is given by

$$V(r) \equiv \begin{cases} \alpha(1-r)^4 - \beta(1-r)^2, & r < 1, \\ 0, & r \geq 1, \end{cases} \quad (B.1)$$

where $\alpha$ and $\beta$ are two constants. It is compared with the standard Lennard-Jones 12-6 potential in Fig. 29. The primary units of length and energy are the cut-off distance $r_c$ at which the potential is valued zero and potential well depth $\varepsilon$ respectively, and are both scaled to unity. Consequently, a single parameter $r_0$, the equilibrium interparticle distance, controls the shape of the potential. The distance $r_0$ is normalized with respect to $r_c$ such that it is always less than unity. In this study, $r_0$ is arbitrarily fixed at 0.85. The two parameters $\alpha$ and $\beta$ that modulate the potential are given by:

$$\alpha = (1 - r_0)^{-1}, \quad \beta = 2(1 - r_0)^{-2}. \quad (B.2)$$

Evidently, the range of short-smooth potential is much smaller than the Lennard-Jones (LJ) 12-6 potential and this
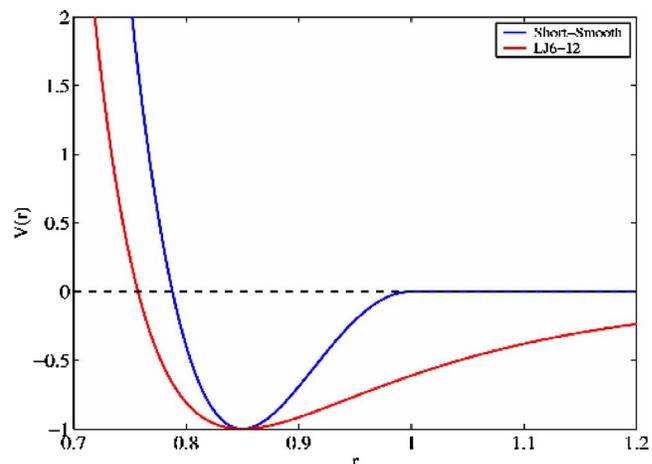


FIG. 29. (Color online) The short-smooth interatomic potential.

translates into significant savings in computational time. Note that the potential energy well ($\varepsilon$) is the same as that of the LJ potential. However the reference length ($\sigma$) for the short-smooth potential is nominally longer than that of the LJ potential.

The phase diagram with the short-smooth potential is unavailable and is expected to be different from that of a LJ system. The liquid gas equilibrium phase shrinks when the attractive part of the potential becomes short ranged [42]. In some sense, short-smooth potential (SS) is a caricature of a conformal, or a truncated and shifted LJ 12-6 potential. Pair potentials are conformal when their plots can be constructed by adjusting the values of $\sigma$ and $\varepsilon$. For two-dimensional (2D) LJ 12-6, the critical temperature and density in reduced units are 0.52 and 0.36, respectively [43]. The critical temperature reduces to 0.46 when the LJ 12-6 potential is truncated at $2.5\sigma$ and shifted. The critical density, however, changes only by a small fraction to 0.35. Similar results are obtained for LJ potentials such as 32-6 that have significantly shorter attractive range [42]. It is clear that the phase dia-
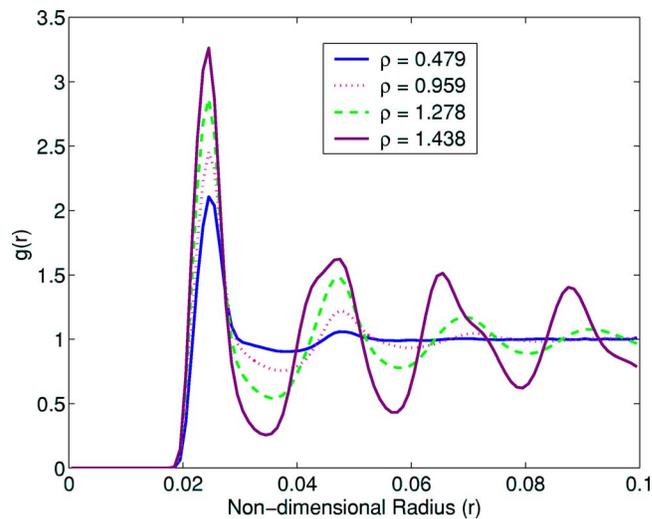


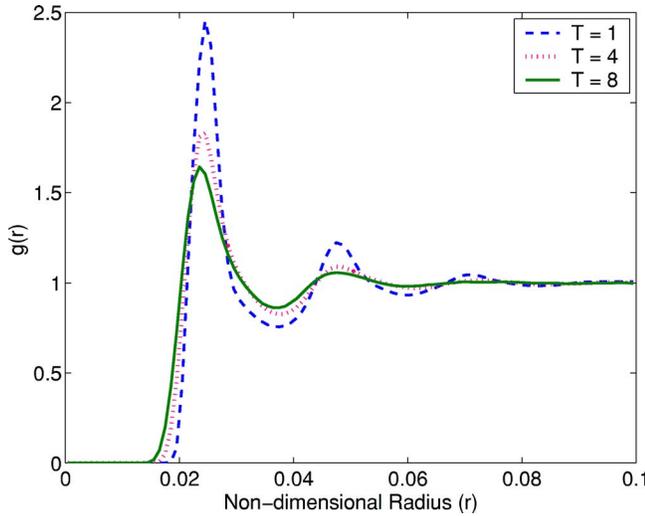FIG. 30. (Color online) The radial distribution function at $T=1$.

FIG. 31. (Color online) The radial distribution function at $\rho$ $=0.96$.

gram with SS potential will be different from that of a 2D LJ 12-6 potential. Based on the similarity with the truncated and shifted potentials, it can be argued that the critical temperature of SS potential will be lower than that with LJ 12-6 potential which is 0.52 in two dimensions.

Radial distribution function (rdf) gives a reasonable confirmation of the thermodynamic state. Figure 30 shows the rdf with short-smooth potential for different densities at a temperature of 1.0. We can note that at $(\rho, T) = (0.48, 1)$, the state of the system is close to that of a gas while it is more solidlike at a density $(\rho, T) = (1.44, 1)$. The effect of temperature on rdf is depicted in Fig. 31. All of the simulations performed in this study are with $(0.48, 1)$.

## APPENDIX C: NUMERICAL SOLUTION OF INCOMPRESSIBLE NAVIER-STOKES EQUATIONS WITH VORTICITY-STREAM FUNCTION APPROACH

The incompressible equations for Newtonian fluid flow in the absence of body forces are given by [36]

$$\boldsymbol{\nabla} \cdot \mathbf{v} = 0, \tag{C.1}$$

$$\rho \left( \frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \boldsymbol{\nabla} \mathbf{v} \right) = - \boldsymbol{\nabla} p + \mu \nabla^2 \mathbf{v}. \tag{C.2}$$

Computational difficulties often associated with the above set of nonlinear equations are alleviated by employing the following mathematical transformation [37]:

$$\boldsymbol{\omega} = \boldsymbol{\nabla} \times \mathbf{v}, \tag{C.3}$$

$$u = \frac{\partial \psi}{\partial y}, \quad v = - \frac{\partial \psi}{\partial x}, \tag{C.4}$$

$\boldsymbol{\omega} = \omega \mathbf{k}$, where $\mathbf{k}$ is the unit vector along the axis perpendicular to the computational domain. $\psi$ is the stream function, and $u$ and $v$ denote the two components of the velocity vec-

tor. Note that this transformation automatically satisfies the mass continuity equation. The transport of vorticity can be written in terms of a single scalar equation [37]. It is given by

$$\frac{\partial \omega}{\partial t} + u \frac{\partial \omega}{\partial x} + v \frac{\partial \omega}{\partial y} = \left( \frac{1}{\mathrm{Re}_L} \right) \nabla^2 \omega. \tag{C.5}$$

As can be noted, vorticity is convected by the velocity components (gradients of the stream function) and also gets dissipated. The vorticity is connected to the streamfunction through the following equation:

$$\nabla^2 \psi = - \omega. \tag{C.6}$$

A significant advantage of the vorticity-streamfucntion approach is in the elimination of the pressure term. This facilitates a solver on regular nonstaggered grids. In addition, there is only one transport equation as against two in the primitive variables approach. The streamfunction equation also falls in the Poisson class which has very efficient numerical solution techniques such as Krylov and multigrid methods.

The streamfunction and vorticity transport equations are solved with a finite-difference algorithm on a nonstaggered grid [37]. At the solid walls, the velocity components are zero by no-slip assumption. Hence, the streamfunction is arbitrarily set to zero at the stationary walls. Since the boundary conditions for vorticity do not exist in a natural way, the streamfunction is expanded with the Taylor's series and a second order finite difference approximation for the vorticity is applied on the walls [44].

We have adopted the direct numerical algorithm given by Pozrikidis [37]. We start with a guessed vorticity field. New streamfunction values are evaluated by solving the Poisson equation for streamfunction, Eq. (C.6). The velocity components are calculated on all the interior nodes and the boundary vorticity values are evaluated. Next the vorticity field is predicted using the steady-state version of Eq. (C.5). The $L_2$ norm of the vorticity is compared between successive iterations. Convergence is established when the relative error becomes less than an arbitrarily assigned value of $10^{-6}$.

## APPENDIX D: GENERALIZED VORTICITY AND STREAMFUNCTION

For steady flow, the mass conservation is:

$$\boldsymbol{\nabla} \cdot (\rho \mathbf{v}) = \frac{\partial(\rho u)}{\partial x} + \frac{\partial(\rho v)}{\partial y} = 0. \tag{D.1}$$

Define a generalized streamfunction $\Omega$ such that

$$\frac{\partial \Omega}{\partial y} = \rho u \text{ and } \frac{\partial \Omega}{\partial x} = - \rho v. \tag{D.2}$$

The function $\Omega$ is similar to the streamfunction $\psi$ for incompressible flow. Now, define a generalized vorticity $\zeta$ as

$$\zeta \equiv \boldsymbol{\nabla} \times (\rho \mathbf{v}). \tag{D.3}$$

The scalar component (for 2D) is given by

$$\zeta = \frac{\partial(\rho v)}{\partial x} - \frac{\partial(\rho u)}{\partial y}. \tag{D.4}$$

Combining Eqs. (D.2) and (D.4) we get:

$$\nabla^2 \Omega = -\zeta. \tag{D.5}$$

Equation (D.5) is similar to the incompressible equation $\nabla^2 \psi = -\omega$, where $\omega$ signifies the standard vorticity. $\Omega$ has the same physical interpretation as that of $\psi$. The difference (in the values) of two streamlines is proportional to the mass flow rate between them.

---

[1] X. B. Nie, S. Y. Chen, W. N. E, and M. O. Robbins, J. Fluid Mech. **500**, 55 (2004).

[2] R. Delgado-Buscalioni and P. V. Coveney, Phys. Rev. E **67**, 046704 (2003).

[3] E. G. Flekkoy, G. Wagner, and J. Feder, Europhys. Lett. **52**, 271 (2000).

[4] J. Li, D. Liao, and S. Yip, J. Comput.-Aided Mater. Des. **6**, 95 (1999).

[5] N. G. Hadjiconstantinou, J. Comput. Phys. **154**, 245 (1999).

[6] N. G. Hadjiconstantinou and A. T. Patera, Int. J. Mod. Phys. C **8**, 967 (1997).

[7] S. T. O'Connell and P. A. Thompson, Phys. Rev. E **52**, R5792 (1995).

[8] I. G. Kevrekidis, C. W. Gear, J. H. Hyman, P. G. Kevrekidis, O. Runborg, and C. Theodoropoulos, Commun. Math. Sci. **1**(4), 715 (2003).

[9] I. G. Kevrekidis, J. Li, and S. Yip, *Multiscale Modelling and Simulation*, edited by S. Attinger and P. Koumoutsakos, Lecture Notes in Computational Science and Engineering, 39 (Springer, Berlin, 2004).

[10] C. W. Gear, J. Li, and I. G. Kevrekidis, Phys. Lett. A **316**, 190 (2003).

[11] J. Li, P. G. Kevrekidis, C. W. Gear, and I. G. Kevrekidis, Multiscale Model. Simul. **1**(3), 391 (2003).

[12] C. Theodoropoulos, Yue. H. Qian, and I. G. Kevrekidis, Proc. Natl. Acad. Sci. U.S.A. **97**, 9840 (2000).

[13] J. Li, D. Liao, and S. Yip, Phys. Rev. E **57**, 7259 (1998).

[14] E. T. Jaynes, Phys. Rev. **106**, 620 (1957); **108**, 171 (1957).

[15] G. R. Liu and M. B. Liu, *Smoothed Particle Hydrodynamics* (World Scientific, Singapore, 2003).

[16] D. C. Rapaport, Phys. Rev. A **36**, 3288 (1987).

[17] D. C. Rapaport, Phys. Rev. Lett. **57**, 695 (1986).

[18] J. Koplik and J. R. Banavar, Phys. Fluids **7**, 3118 (1995).

[19] J. Koplik and J. R. Banavar, J. Rheol. **41**, 787 (1997).

[20] E. Parzen, Ann. Math. Stat. **33**, 1065 (1962).

[21] B. W. Silverman, *Density Estimation for Statistics and Data Analysis* (Chapman & Hall, London, 1994).

[22] A. Mohammad-Djafari, *Maximum Entropy and Bayesian Methods*, Proceeding of the Eleventh International Workshop on Maximum Entropy Methods of Statistical Analysis, edited by C. R. Smith, G. J. Erickson, and P. O. Neudorfer (Kluwer Academic Publishers, Netherlands, 1992).

[23] C. E. Shannon, Bell Syst. Tech. J. **27**, 379 (1948).

[24] J. N. Kapur, *Maximum Entropy Models in Science and Engineering* (Wiley Eastern Ltd., New Delhi, 1989).

[25] D. J. Evans and G. P. Morriss, *Statistical Mechanics of Nonequilibrium Liquids* (Academic Press, San Diego, 1990).

[26] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C++* (Cambridge University Press, Cambridge, 2002).

[27] G. Talenti, Inverse Probl. **3**, 501 (1987).

[28] P. N. Shankar and M. D. Deshpande, Annu. Rev. Fluid Mech. **32**, 93 (2000).

[29] K. J. Van Vliet, J. Li, T. Zhu, S. Yip, and S. Suresh, Phys. Rev. B **67**, 104105 (2003).

[30] D. C. Rapaport, *The Art of Molecular Dynamics Simulation*, 2nd ed. (Cambridge University Press, Cambridge, 2004).

[31] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Clarendon Press, Oxford, 2003).

[32] Ju Li, Modell. Simul. Mater. Sci. Eng. **11**, 173 (2003).

[33] S. Y. Liem, D. Brown, and J. H. R. Clarke, Phys. Rev. A **45**, 3706 (1992).

[34] C. Y. Soong, S. H. Wang, and P. Y. Tzeng, Phys. Fluids **16**, 2814 (2004).

[35] J. P. Boon and S. Yip, *Molecular Hydrodynamics* (Dover Publication, Inc., New York, 1991).

[36] R. Aris, *Vectors, Tensors and the Basic Equations of Fluid Mechanics* (Dover, New York, 1989).

[37] C. Pozrikidis, *Introduction to Theoretical and Computational Fluid Dynamics* (Oxford University Press, New York, 1997).

[38] *Fluent 6.1 User's Guide* (Fuent Inc., Lebanon, New Hampshire, 2001).

[39] X. Fan, N. Phan-Thien, N. T. Yong, and X. Diao, Phys. Fluids **14**, 1146 (2002).

[40] A. Okabe, B. Boots, K. Sugihara, and S. N. Chiu, *Spatial Tessellations* (John Wiley and Sons, New York, 2000).

[41] D. W. Scott, *Multivariate Density Estimation: Theory, Practice, and Visualization* (John Wiley & Sons, New York, 1992).

[42] H. Okumura and F. Yonezawa, J. Chem. Phys. **113**, 9162 (2000).

[43] B. Smit and D. Frenkel, J. Chem. Phys. **94**, 5663 (1991).

[44] J. C. Tannehill, D. A. Anderson, and R. H. Pletcher, *Computational Fluid Mechanics and Heat Transfer* (Taylor and Francis, Washington, D.C, 1997).