Coupling continuum to molecular-dynamics simulation: Reflecting particle method and the field estimator

Ju Li, Dongyi Liao, and Sidney Yip
Department of Nuclear Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
(Received 11 November 1997)

We propose a method to induce microchannel (Poiseuille) flow in a molecular-dynamics simulation by introducing a partially reflecting ''membrane'' as a means of driving the fluid flow while conserving particle number and total energy. We also develop a method to estimate various continuous macroscopic fields from particle data, based on maximum likelihood inference. A general statistical approach is discussed for coupling the continuum with molecular-dynamics simulation with emphasis on minimal disturbance to particle dynamics, which is to be fully developed later. [S1063-651X(98)15105-X]

PACS number(s): 02.70.Ns, 47.11.+j, 83.10.Pp

I. INTRODUCTION

In recent years there has been increasing interest in simulating dynamical phenomena in statistical systems that are manifest on different length scales. In a fluid flow problem, one can imagine decomposing the domain of interest into a relatively confined region, to be treated in atomistic detail, and a larger, remainder region where the continuum description is appropriate. The fundamental question then is how to optimally couple the atomistic region to the continuum [1,2]. From the standpoint of computations at either the microscopic or the macroscopic level, a central issue is how to determine and impose the proper boundary conditions in an internally consistent manner.

In this two-paper sequence we describe an approach to deal with this problem, which in several aspects can be shown to be mathematically optimal. In the first paper we concentrate on two methodological developments, one is a method to drive fluid flow for use in molecular-dynamics (MD) simulations, the other is a method to estimate continuous macroscopic fields of density, temperature, and velocity $\{\rho(\mathbf{x}),T(\mathbf{x}),\overline{\mathbf{v}(\mathbf{x})}\}$ using the discrete particle data given by an MD simulation. While we demonstrate that each method is useful for its own purpose, they can be combined in a general scheme for coupling continuum and atomistic simulations. This is the subject of the second paper [3] where we develop a procedure for imposing desired thermodynamic field boundary conditions in an atomistic simulation, utilizing the information given by the field estimator; it can be formulated as a transformation that is optimal in the sense that it will minimize the artificial disturbances to particle dynamics.

II. REFLECTING PARTICLE METHOD (RPM)

Consider Poiseuille flow in a narrow channel where compressible fluid enters and leaves under a pressure difference $P_{\text{in}}-P_{\text{out}}>0$. In the channel, body forces will be neglected. As the channel width decreases to μm range, it is expected that significant variations in density and temperature will occur such that a nonlinear pressure drop develops along the channel [4], a behavior that is different from the simple prediction of incompressible Navier-Stokes equation. Further complications arise due to the fluid-solid interactions at the interface, where simple boundary conditions such as the assumption of ''no-slip'' may become inadequate [5]. To probe such fine-scale phenomena MD simulation [6] can be used; since it is often neither possible nor necessary to treat the entire problem using discrete particles, a method to couple the MD region to the surrounding continuum would be needed.

In a standard continuum description, the Navier-Stokes equation for incompressible Newtonian fluid [7], we have

$$\rho\left(\frac{\partial\mathbf{v}}{\partial t}+\mathbf{v}\cdot\nabla\mathbf{v}\right)=-\nabla P+\mu\nabla^2\mathbf{v}+\rho\mathbf{g}, \tag{2.1}$$

where μ is the shear viscosity. For steady state flow: $\partial\mathbf{v}/\partial t=0$, and ignoring gravitational body force, the forces that act on the fluid come from the pressure gradient term $-\nabla P$ and the viscous term $\mu\nabla^2\mathbf{v}$, which means that the fluid under consideration should only be pushed by other fluids.

To simulate fluid flow in MD we imagine a *region of interest* $\mathcal{C}$ where the fluid atoms are not subjected to artificial forces or constraints, and another region $\mathcal{A}$ sufficiently far away from $\mathcal{C}$ where actions are applied to the particles to achieve the desired flow conditions. Because of molecular chaos, perturbations to particle dynamics will decay over a distance of a few mean free paths, thus setting a lower limit for the $\mathcal{C}$-$\mathcal{A}$ separation. What we have in mind, in essence, is a feedback control system where actions in $\mathcal{A}$ cause the prescribed flow conditions to be induced in $\mathcal{C}$, while particles in $\mathcal{C}$ still follow their natural dynamics. In fact, a fluid atom in $\mathcal{C}$ should not be able to distinguish any difference with reality, and how the flow is induced should be irrelevant (for us, of course, there is the problem of cost effectiveness).

Current methods simulating the Poiseuille flow seem to fall into two types, the first may be called the ''gravitation method,'' [8] where a constant artificial acceleration field is imposed on all the atoms of the fluid. In our opinion there are three unsatisfactory aspects to this approach. (1) In order to induce appreciable flow the acceleration field would need to be as high as 10^{12} times the earth's gravitation. As a consequence, significant kinetic-energy rescaling is required to re-

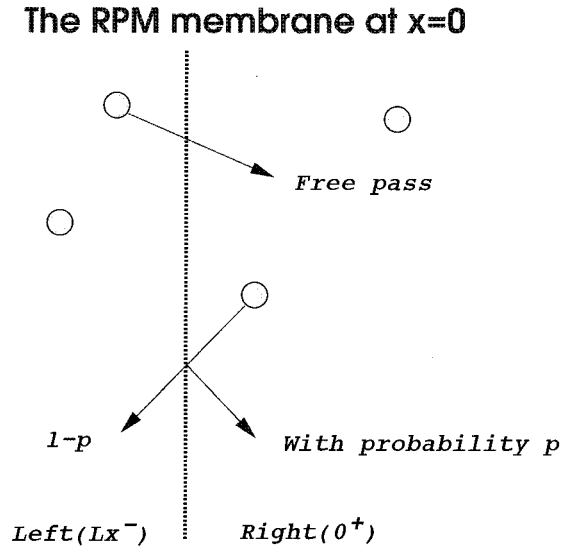

FIG. 1. The reflecting particle method. If a fluid atom crosses $x=0$ from left ($x=L_X^-$) to right ($x=0^+$), it passes through 100%; but if it crosses $x=0$ from right to left, then it could be elasticly reflected with probability p. The membrane is otherwise transparent in the sense that particles on two sides can interact.

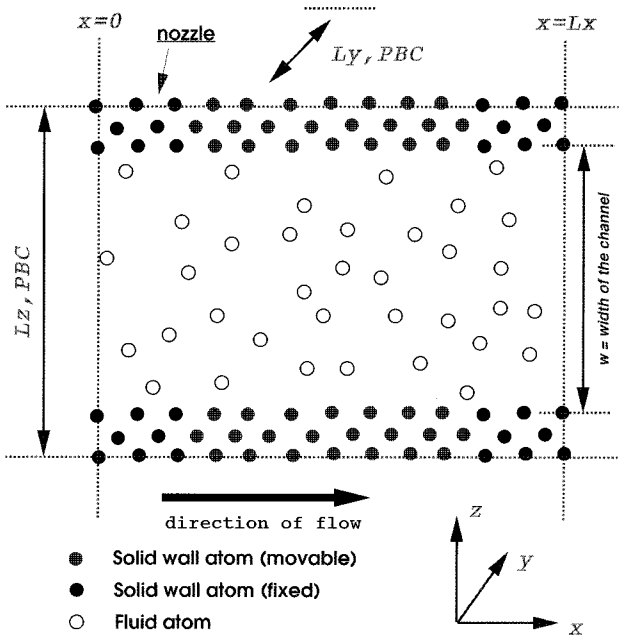

FIG. 2. Setup of the simulation cell. $x=0^+$ and $x=L_X^-$ are two sides of the RPM membrane (see Fig. 1), which drives the fluid flow. The nozzle consists of two units of fixed fcc solid wall atoms because pressure variation near the membrane is big; other wall atoms are movable. Between the nozzles the correct particle dynamics can be assumed to be restored, thus becoming our region of interest, $\mathcal{C}$.

move the heat generated. (2) In the continuum limit one does not recover the Navier-Stokes equation (2.1) unless the pressure gradient $-\nabla P$ is assumed to be constant everywhere. (3) The effects on local dynamics induced by constant acceleration are quite different from those induced by $-\nabla P$, which have microscopic spatial and temporal fluctuations associated with molecular collisions. Besides these drawbacks, the rescaling of kinetic energy to remove heat will also disturb the particle dynamics.

The second approach to simulating Poiseuille flow may be called the ''reservoir method'' [9], where artificial particle source and sink, at the two ends of the simulation cell, are maintained at different pressures. This approach is similar in spirit to our method; however, there are differences in implementation. In the existing method, in order to maintain the inlet reservoir, the source region is compressed every few hundred steps and new particles are injected into the empty space. The source reservoir takes about 1/4 of the entire simulation cell which amounts to a considerable portion of the computational burden. Also, the compression-injection procedure may not give a very smooth and steady flow.

Our aim is to simulate fluid flow inside a microchannel with minimal unnatural perturbations or constraints. It is well known that this process, whether in steady state or not, is dissipative in that the work to drive the fluid flow is continuously transformed by viscous action into heat, which then must be removed from the simulation cell. Other requirements which one could demand in setting up the pressure difference between inlet and outlet are current conservation (such that steady state is attainable), easy implementation, and smooth approach to steady state.

We propose a simple method for driving fluid flow which satisfies these requirements (see Fig. 1). The method consists of introducing a fictitious membrane to act as a filter to allow atoms crossing from one direction to pass through without hindrance, while atoms crossing from the other direction are elastically reflected with a certain probability p. The membrane has no other effects on the atoms; atoms interact with each other across the membrane in the same manner as anywhere else in the fluid.

It is intuitive that in the present setup the fluid would be driven in the direction of free pass (call this the x direction, and let the membrane be located at $x=0$), and a negative pressure gradient ($\Delta P/\Delta X<0$) would be established. It is also clear that since no new particle or energy is injected at any time, the particle number and total energy of the system are conserved, thus allowing steady state flow to be possible. Moreover, one does not encounter situations where particles find themselves in energetically unfavorable positions, which can occur in methods using random particle insertion. It will be seen below that our method does not require extra kinetic-energy rescaling, the total energy being as well conserved as in any bulk liquid simulation.

To illustrate our method we consider three-dimensional microchannel flow (see Fig. 2) of liquids in which the atoms interact through a Lennard-Jones 6-12 potential with radial cutoff at 2.2σ [5]. Periodic boundary conditions are imposed in all three directions. Notice that translational invariance does not hold in the x direction because of the membrane at $x=0$. Along the z direction the channel, with width w, is bounded by two parallel solid walls consisting of (001) layers of fcc planes with atoms interacting via a stronger Lennard-Jones potential, same σ but with $\epsilon_{ww}=5\epsilon_{wf}=5\epsilon_{ff}$, such that at the fluid temperature it is in solid phase. Also, the wall atoms are given a larger mass ($m_w=5m_f$). The choice of these parameters leads to an intrinsic wall length/time scale comparable to that of the fluid, thus enhancing fluid-wall coupling. The first two units of the wall atoms near $x=0$ are fixed (to serve as nozzles) because pres-

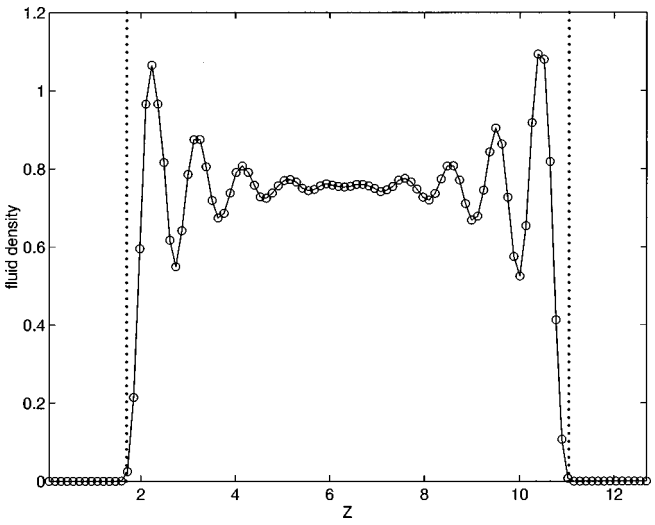

FIG. 3. Average cross-sectional (Z) fluid density. The vertical dotted lines indicate the theoretical positions of fluid-solid interfaces. Empty circles are results by dividing the entire cell into 100×100 bins in xz, and average over x in the region of interest (between the two nozzles). The solid line is its interpolation.

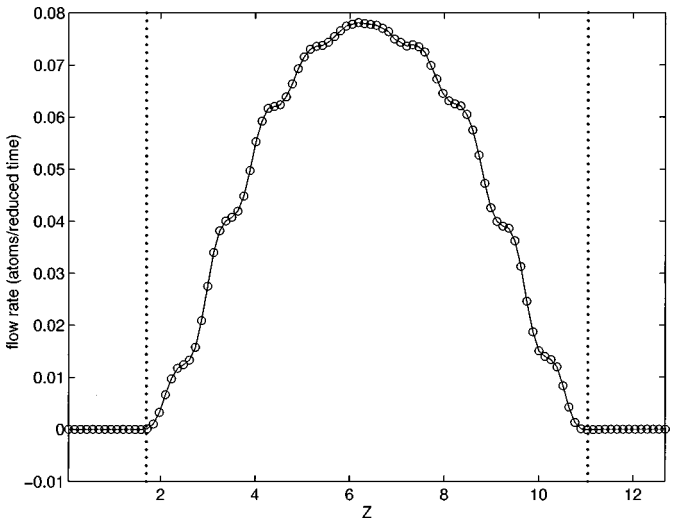

FIG. 4. Average cross-sectional fluid flow rate. The way data are measured is explained in Fig. 3.

sure variation near the membrane could be large, and to prevent mechanical failure at other parts of the channel we attach fourth-order springs to the wall atoms, which under normal conditions should not influence the equilibrium lattice dynamics of the solid.

$O(N)$ efficiency can be achieved in the calculation by dividing the simulation cell into a number of bins (exploiting the short-range nature of the interactions). Each bin maintains a list (the bin list) of all the particles inside and each particle maintains a neighbor list, whose range is slightly larger than the potential cutoff. A ''flash'' condition is devised for updating the neighbor lists of atoms. The neighbor lists make force evaluations $O(N)$, and the bin lists make updating the neighbor lists $O(N)$ (because updating is restricted to atoms in the nearby 27 bins if the bin size is greater than the neighbor list range). Data dependency between bins includes force evaluation, neighbor list maintainance, and particle transfer. The code has been implemented on both single processor workstations and on parallel shared memory SMP's with excellent speed scaling.

In order to compare with literature results [5], we have studied Poiseuille flow with average reduced temperature 1.1 and reduced density 0.81 for both fluid and wall [10]. The simulation cell is $51.0\times5.1\times12.7$ in the xyz directions and contains 1783 fluid atoms and 900 solid wall atoms (two fcc layers). The run length is 5×10^5 time steps (step size $=0.005$). The reflection probability p is chosen to be 1. After equilibrating for 40 000 steps, the RPM is activated at $x=0$ for another 40 000 steps before starting to accumulate properties in 100×100 meshed bins along xz. Flow rates are measured to be 3.093 ± 0.007 atoms/reduced time at many

different x's; the small variation along x indicates that steady state is well achieved.

The average cross-sectional fluid density is shown in Fig. 3, where the vertical dotted lines indicate the positions of the fluid-solid interfaces derived from the solid density [lattice parameter$=(4/\rho)^{1/3}\approx1.7$].

As the fluid density varies significantly near the wall, it is more helpful to look at the flow rate ($\propto\rho v_x$) rather than v_x. The average cross-sectional flow rate is shown in Fig. 4. One can see that the overall shape agrees well with the parabolic profile predicted by the incompressible Navier-Stokes equation, and is smoother than either the particle density or the velocity profile. Microscopic details such as layering near the solid walls can still be distinguished.

The stress distribution ($\tau_{ij}\equiv-P_{ij}$) throughout the channel is calculated by attributing pair interaction terms to pair centers, and ''kinetic energy'' terms to particle sites. It is found that the diagonal elements of the stress tensor (P_{xx}, P_{yy}, P_{zz}) are reasonably close to their average in the fluid stream. The average cross-sectional shear stress profile is shown in Fig. 5, one sees that near the stream center P_{xz} is linear with z, in agreement with the continuum solution.

Midstream P_{xx} (averaged across the middle 15 bins) versus x are shown in Fig. 6, where the vertical dotted lines indicate positions of the two fixed fcc wall units near the RPM membrane. In between the dotted lines we regard the particle dynamics to be essentially restored, and it becomes

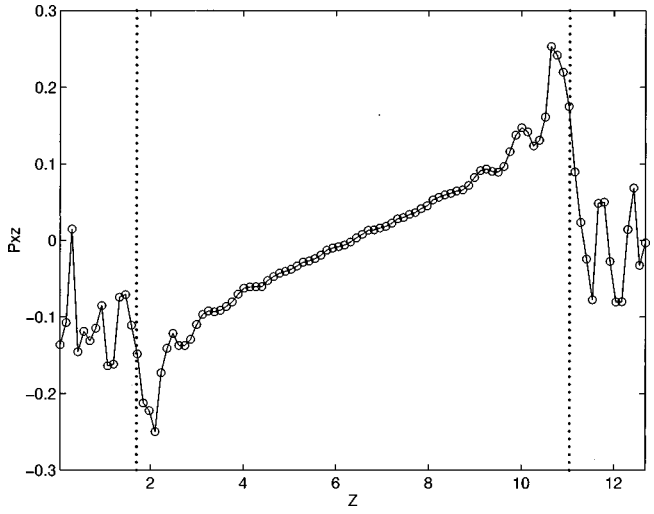

FIG. 5. Average cross-sectional local shear stress ($P_{xz}\equiv-\tau_{xz}$). The way data are measured is explained in Fig. 3.

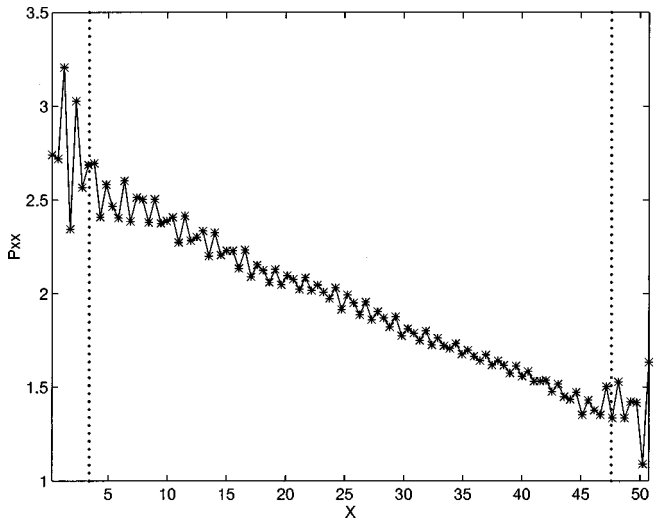

FIG. 6. Midstream pressure (P_{xx}) with respect to x. The vertical dotted lines indicate the positions of the nozzles. Stars are results obtained by dividing the entire cell into 100×100 bins in xz, and averaging over 15 bins near $z=L_Z/2$.

our region of interest, $\mathcal{C}$. The main objective of this method, that of building up a pressure difference between two ends of the channel, is seen to be realized, and its linear decay along the channel in $\mathcal{C}$ is in agreement with the simple continuum solution. Over the entire region temperature variation is less than 5% and density variation is less than 6%, thus in this illustrative example fluid properties are not significantly altered along the flow to produce a nonlinear pressure drop. We expect that with increasing channel length, one will be able to observe nonlinear behavior, signalling that the simple Navier-Stokes solution can no longer describe the entire channel.

For a quantitative comparison with the continuum description (2.1), let us calculate the fluid shear viscosity μ, using the measured total flow rate J and the pressure gradient $-(\Delta P/\Delta x)$ achieved in this simulation. Assuming Eq. (2.1) is correct, the velocity field should be

$$v_x(z)=\frac{1}{2\mu}\left(-\frac{\Delta P}{\Delta x}\right)\left[\left(\frac{w}{2}\right)^2-z^2\right], \tag{2.2}$$

so the total flow rate should be

$$J=\rho L_Y\int_{-w/2}^{w/2}\frac{1}{2\mu}\left(-\frac{\Delta P}{\Delta x}\right)\left[\left(\frac{w}{2}\right)^2-z^2\right]dz=\frac{-\rho\Delta P L_Y w^3}{12\Delta x\mu}. \tag{2.3}$$

Taking $J=3.093$, $\rho=0.81$, $L_Y=5.1$, $w=L_Z-2(4/\rho)^{1/3}=9.29$, and obtaining $-(\Delta P/\Delta x)=0.027\,65$ by least-square fitting of Fig. 6 inside $\mathcal{C}$, we get $\mu_{\rm sim}=2.47$, which is in fair agreement with the value $\mu=2.14$ obtained entirely independently from a bulk liquid simulation [1].

As a subtle feature of the RPM method proposed here, we find that the heat generated by viscous action inside the fluid and by possible fluid-solid slip at the interface [11] is actually removed by constant entropy extraction through the action of the membrane, which in effect constitutes a ''Maxwell's demon'' system [12]. To quantify this claim we propose the following heat balance equation:

$$\begin{aligned}k_BT(&\ln C_{f_1+f_2}^{f_1}-\ln C_{f_1+f_2}^{f_1+pf_2})\\&=k_BT\left[f_1\ln\left(1+\frac{pf_2}{f_1}\right)+pf_2\ln\left(p+\frac{f_1}{f_2}\right)\right.\\&\quad\left.+f_2(1-p)\ln(1-p)\right]\\&\approx\int d\Omega\,\tau_{ij}v_{i,j}+(\text{interfacial friction}),\end{aligned} \tag{2.4}$$

where $f_1(f_2)$ is the rate of particles hitting $x=0$ from left (right) when the system reaches steady state, based on an estimation of the entropy ($S=k_B\ln\Omega$) extraction rate. The rationale is that *if* the membrane does not act, then in a unit period of time, the macroscopic parameters specifying the system are f_1 and f_2, which corresponds to $C_{f_1+f_2}^{f_1}=(f_1+f_2)!/f_1!f_2!$ microstates (ways of crossing); after the membrane interferes it is changed to $C_{f_1+f_2}^{f_1+pf_2}$, as the system cannot distinguish between atoms that hit from the right and get reflected and atoms that originally hit from the left [13]. The first term on the right-hand side (RHS) is the dissipation kernel (heat generation rate) integrated inside the fluid bulk [7], assuming the continuum approach to be applicable. The second term, due to interfacial friction, is more involved but we expect it to be small when there is no ''stick-slip'' motion [11]. The left-hand side (LHS) and RHS are equated by the fundamental relation $T\Delta S=\Delta Q$.

Overall we should have the number balance equation,

$$f_1-(1-p)f_2=J, \tag{2.5}$$

to hold at steady state, where J is the total flow rate.

Equation (2.4) has been explicitly verified by simulation for many cases and is found to hold quantitatively in most situations. Assuming the continuum treatment to be valid, we can write for the heat generated by viscous dissipation inside the fluid [7],

$$\frac{dQ}{dt}=\int d\Omega\,\tau_{ij}v_{i,j}=\int d\Omega\,\mu\left(\frac{dv_x}{dz}\right)^2=\frac{L_XL_Yw^3}{12\mu}\left(\frac{\Delta P}{\Delta x}\right)^2, \tag{2.6}$$

which is to be balanced with the entropy extraction rate at the membrane.

Table I summarizes all the simulation results for the above $51.0\times5.1\times12.7$ system with different reflection probability p's. Depending on the magnitude of p, the simulation run length varies from 5×10^5 to 3×10^6 timesteps until all relevant quantities have fully converged. $\mu_{\rm sim}$ is then calculated using Eq. (2.3), and is used to evaluate Eq. (2.6).

Figure 7 compares the LHS and RHS of Eq. (2.4) with varying p's. The agreement, when $0.1<p<0.8$, is within 15% as the dissipation rate varies by two decades in the range. If we take into consideration all the possible errors and ambiguities involved in evaluating Eq. (2.6), such as that $\rho_{\rm midstream}$ is not exactly ρ and w may plausibly take values other than $L_Z-2(4/\rho)^{1/3}$, this is remarkably good agreement. In particular, the parts of the fluid that are near the membrane (see Fig. 6) can not be taken as normal fluid, in contrast to what is assumed in deriving Eq. (2.6), and they tend to be

TABLE I. Verification of the heat balance equation (2.4).

p	0.050	0.100	0.150	0.1875	0.225	0.300	0.400	0.500	0.600	0.700	0.850	1.000
f_1	18.2322	17.7521	17.2971	16.9136	16.5921	15.8123	14.5614	13.3660	11.8653	10.2939	7.1772	3.0934
f_2	19.0759	19.5294	20.0876	20.3828	20.8076	21.6837	22.9158	24.4480	26.1611	28.3762	32.3390	38.0656
$-\Delta P/\Delta X$	0.00068	0.00144	0.00227	0.00313	0.00420	0.00544	0.00744	0.01006	0.01285	0.01564	0.02115	0.02765
$\mu_{\rm sim}$	1.8291	2.0793	2.8100	2.4567	2.5776	2.3700	2.4689	2.5100	2.5100	2.4000	2.4900	2.4700
LHS	0.0056	0.0184	0.0360	0.0725	0.1173	0.2220	0.4055	0.7660	1.2306	2.0246	4.0097	10.9804
RHS	0.0044	0.0174	0.0318	0.0696	0.1188	0.2170	0.3904	0.7020	1.1423	1.7734	3.1262	5.3843

more ''turbulent'' than fluids in $\mathcal{C}$ and thus should generate more heat. So the observation that the LHS of Eq. (2.4) is consistently a little higher than the RHS seems to make sense.

A much bigger system ($76.5\times10.2\times20.4$) with 10 179 fluid atoms and 2700 wall atoms has been studied in a 1 000 000 time step run with $p=0.3$, under the same temperature and density conditions. All the qualitative features of our discussion above remain valid although the system is now closer to a continuum. We find that $f_1=56.8445$, $f_2=73.8267$, $-(\Delta P/\Delta x)=0.003\,624$, and the inferred $\mu_{\rm sim}=2.37$. In the heat balance equation, the LHS gives 1.7692 while the RHS gives 1.7681.

III. THERMODYNAMIC FIELD ESTIMATOR

The problem we raise is the following. Given a set of particle data (we assume the particles to be of the same type, with reduced mass 1) $\{(\mathbf{x}_i,\mathbf{v}_i),i=1,\ldots,N\}$, where N is a large number, how does one determine the corresponding spatial distributions of fields $\{\rho(\mathbf{x}),T(\mathbf{x}),\overline{\mathbf{v}}(\mathbf{x})\}$, which have meaning in the continuum approach. A conventional way to proceed would be to divide the cell into bins, and average over each bin:

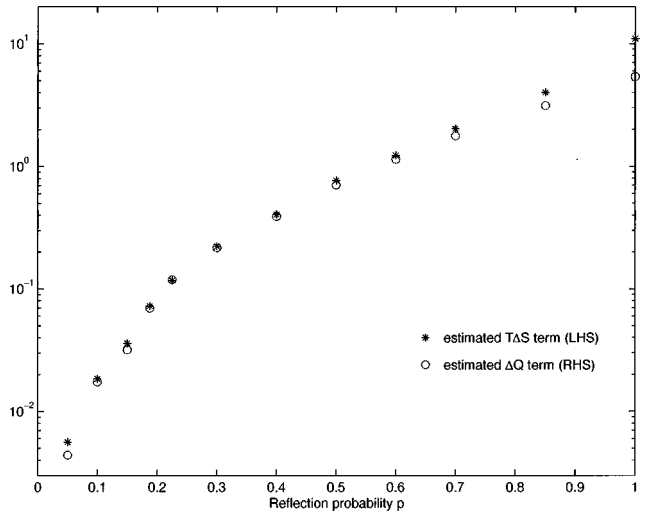

FIG. 7. Simulation verification of the heat balance equation (2.4), with different reflection probability p's. The LHS and RHS of Eq. (2.4) are both theoretical predictions of the system dissipation rate under steady-state flow, one based on the proposed entropy extraction rate from statistical mechanics, the other from the continuum mechanics dissipation kernel.

$$\rho_\alpha=\Omega_\alpha^{-1}\sum_{i\in\text{bin}\alpha}1,$$

$$\overline{\mathbf{v}}_\alpha=\langle\mathbf{v}_i\rangle_{i\in\text{bin}\alpha},\tag{3.1}$$

$$T_\alpha=D^{-1}\langle|\mathbf{v}_i-\overline{\mathbf{v}}_\alpha|^2\rangle_{i\in\text{bin}\alpha}.$$

The set of values $\{\rho_\alpha,T_\alpha,\overline{\mathbf{v}}_\alpha\}$ will represent a spatial distribution.

However, one could ask whether Eq. (3.1) is the best one can do, or the given particle information has been optimally utilized. For the purpose of obtaining continuous streamlines of a flow field, it may turn out that the bin-averaged results $\{\rho_\alpha,T_\alpha,\overline{\mathbf{v}}_\alpha\}$ are not sufficiently smooth so that additional interpolation is necessary. Also, in dividing the cell into bins one may be forced to compromise between spatial resolution and statistical accuracy, which often does not lead to a satisfactory solution.

Suppose we assume that the particles of interest conform to a local Maxwellian distribution,

$$\begin{aligned}dP&=f_M(\mathbf{x},\mathbf{v}|\{\rho(\mathbf{x}),T(\mathbf{x}),\overline{\mathbf{v}}(\mathbf{x})\})d\mathbf{x}d\mathbf{v}\\&=\rho(\mathbf{x})d\mathbf{x}\cdot\frac{1}{[2\pi T(\mathbf{x})]^{D/2}}\exp\left(-\frac{|\mathbf{v}-\overline{\mathbf{v}}(\mathbf{x})|^2}{2T(\mathbf{x})}\right)d\mathbf{v},\end{aligned}\tag{3.2}$$

where D is the dimensionality of the system and we have taken $k_B=1$. This assumption can be explicitly checked whenever necessary; theoretically it should hold quite well when the field gradients are small and the transients have died away. We now propose a method that directly gives continuous field distributions from particle data, based on the concept of maximum likelihood inference [14] in statistics. One picks a spatial basis set: $\{T_l(\mathbf{x}),l=0,\ldots,L-1\}$, typically low-order polynomials, and expresses the field(s), say $\beta(\mathbf{x})=1/T(\mathbf{x})$, as

$$\beta(\mathbf{x})=\sum_{l=0}^{L-1}b_lT_l(\mathbf{x}).\tag{3.3}$$

One then maximizes the total probability P, given the set of data points $\{(\mathbf{x}_i,\mathbf{v}_i),i=1,\ldots,N\}$,

$$P(\{b_l\})=\prod_{i=1}^{N}f_M(\mathbf{x}_i,\mathbf{v}_i|\{\rho(\mathbf{x}),T(\mathbf{x})=\beta^{-1}(\mathbf{x}),\mathbf{v}(\mathbf{x})\})\tag{3.4}$$

with respect to the coefficients $\{b_l\}, l=0,\ldots,L-1$. The maximum is then regarded as the most ''probable'' field distribution in parameter space $\{b_l\}$. In this case we assumed the other two fields $\rho(\mathbf{x})$ and $\mathbf{v}(\mathbf{x})$ to be known [15].

In general the fields should be treated altogether. From Eqs. (3.2) and (3.4) we see that

$$\ln P = \sum_{i=1}^{N} \left(\frac{D}{2} \ln\beta(\mathbf{x}_i) - \frac{\beta(\mathbf{x}_i)|\mathbf{v}_i - \overline{\mathbf{v}}(\mathbf{x}_i)|^2}{2} \right) + \sum_{i=1}^{N} \ln\rho(\mathbf{x}_i) + \text{const.} \tag{3.5}$$

So the density field is decoupled from $\{T(\mathbf{x}), \overline{\mathbf{v}}(\mathbf{x})\}$ and can be treated separately. From now on we will consider only $T(\mathbf{x}), \overline{\mathbf{v}}(\mathbf{x})$ fields, which have $1+D$ functional degrees of freedom.

We have implemented the scheme using Chebyshev polynomials, which have better numerical stability than $\{x^n\}$'s of the same order [16], as the basis set. For $D>1$ dimensional system one can use their products as basis functions, such as

$$T_{lmn}(\mathbf{x}) = T_l(x) T_m(y) T_n(z),$$

if the field has 3D variation.

By noticing that $\overline{\mathbf{v}}(\mathbf{x})$ appears in Eq. (3.5) only in a squared form, we used a two-step relaxation procedure rather than minimizing all the coefficients at once. Given $\beta(\mathbf{x})$, the field $\overline{\mathbf{v}}(\mathbf{x})$ was relaxed by solving a linear set of equations using the Cholesky decomposition [16] method, in the same manner as in the least square fitting problem. This defines a function operationally that only depends on $\beta(\mathbf{x})$. We then relaxed $\beta(\mathbf{x})$ using the Polak-Ribiere conjugate gradient minimization algorithm, with force function evaluated by numerical differentiation.

To demonstrate the power of this method, we consider a simple liquid in a two-dimensional system that is heated on one side and cooled on the other, with the simulation cell mirror-reflected to form a periodic structure. An instantaneous snapshot of the system is taken (Fig. 8) where the dots represent discrete particle positions (x) and kinetic energies (y). A sliding bin average with bin size 12% of the entire cell was performed to obtain the zig-zag curve as a representation of the instantaneous temperature profile. As can be seen this result is strongly fluctuating. To reduce the fluctuations one could increase the bin size at the cost of losing spatial details.

Using the proposed temperature field estimator with 1D spatial dependence ($L=4$), we obtain the smooth curve shown in Fig. 8. It is seen that while the thermal noise is effectively suppressed, the continuous representation goes through the bin-averaged curve in a very reasonable manner. The physical origin of the method, rather than some *ad hoc* smoothing criteria, lends confidence in the meaningfulness of the estimation, in the sense of optimal inference from the data given. Because the estimator is insensitive to thermal noises, we expect the time evolution of the temperature profile to be slowly varying until it eventually becomes a straight line. The fact that small deviations from a linear profile are expected justifies the use of low-order polynomials as a basis set, which is equivalent to choosing ''priors'' in statistical inference with known knowledge.

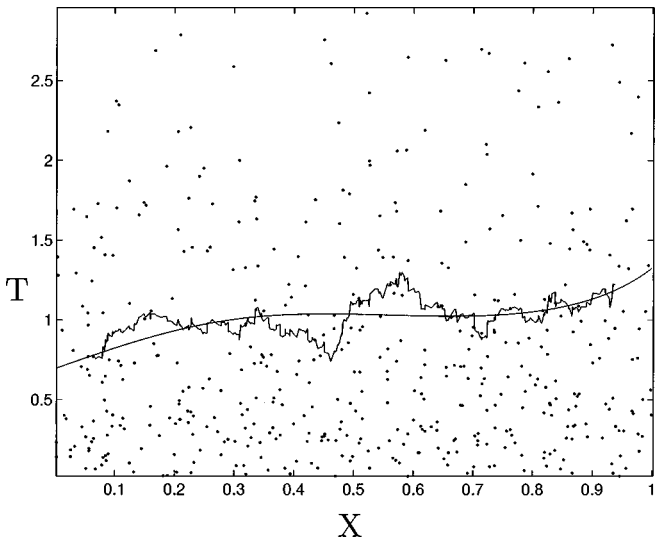

FIG. 8. Snapshot of a 2D liquid system with 500 atoms that is heated on one side ($x=1$) and cooled on the other ($x=0$), where dots represent discrete particle positions (x) and kinetic energies (y). A sliding bin average with size = 0.12 was taken to give the zig-zag curve, representing an instantaneous temperature profile; our method ($L=4$) gives the smooth curve.

Using low-order polynomials as the basis set to represent the fields is the crucial difference when comparing the present method with bin averaging, Eq. (3.1). The latter can be shown to be equivalent to using nonoverlapping staircase functions (idealized as δ functions) as the basis set in the same maximum likelihood inference formalism. Thus, in contrast to representing the temperature profile in Fig. 8 using three or four polynomials, the bin averaging represents the same field using hundreds of ''δ functions,'' which completely ignores any spatial coherence of the data between adjacent bins. We know that the true temperature profile must be a continuous and smooth function irrespective of the thermal noises, nor can it be of strange shape under such a simple setup. The utilization of this spatial coherence knowledge is the reason why polynomial basis sets often exceed bin averaging in performance, especially when the fields are smooth in a regularly shaped domain; on the other hand, under special circumstances, a localized basis may well be helpful when used concurrently with a delocalized basis, in analogy with the situation in electronic structure calculations.

For a specific problem the number of polynomials one should use as basis should correspond to the largest possible field variations one expects to encounter, to the same order of its Taylor series expansion, this field can be satisfactorily cut off. In practice, however, one needs to balance truncation error with statistical error, i.e., a smaller basis set should be used when the data are scant, otherwise thermal fluctuations could dominate the result. For an optimal choice of polynomial order, a detailed mathematical analysis is needed. But a trial and error procedure could be the following. One first proposes a smooth field that has approximately the same magnitude and variations as the field to be estimated. Then one generates artificially a set of ''particle data'' according

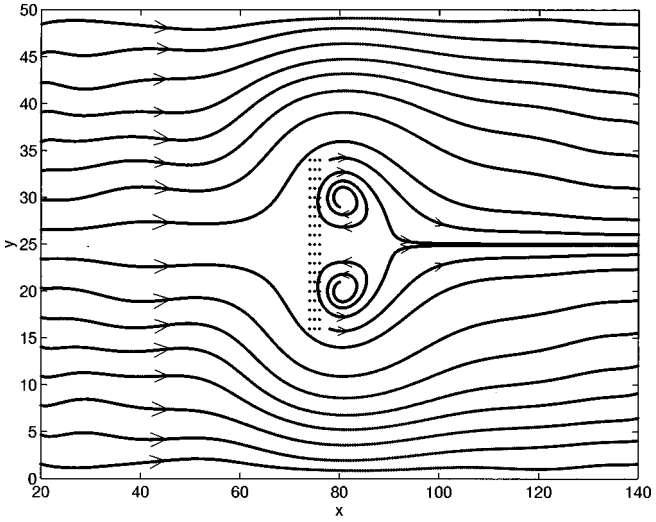

FIG. 9. Streamline plot of 2D fluid flow driven by RPM, with arrows indicating the flow direction. The continuous velocity field is given by the thermodynamic field estimator, based on particle data (system size = 2000 atoms) in 50 000 time steps during an MD simulation. The cell size is (0,0)–(150,50), with a wall barrier of fixed atoms (denoted by dots) in the midstream. The fields are estimated inside a rectangular domain of (20,0)–(140,50) using maximum likelihood inference, by expansion in a continuous and slowly varying basis set.

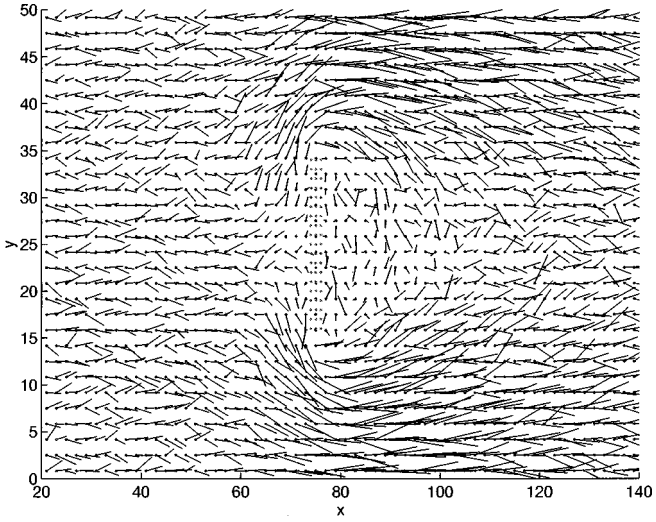

FIG. 10. Corresponding bin-averaged (60×30 mesh) plot of the velocity field, using the same set of particle data as in plotting Fig. 9. The dots indicate the centers of bins and velocities are drawn from them after scaled by 5.

to the proposed field and Eq. (3.2), whose size is the same as the actual data set. Lastly, one feeds the data into the estimator. After repeating the process for a number of times, the order that gives the best average resemblance to the proposed field is the optimum choice.

As another illustration of our method we show (Fig. 9) the streamline plot of fluids flowing over a solid obstacle lying perpendicular to the flow direction. The flow is driven by the RPM. To make everything simple we use a 2D system (150×50, 2000 particles, $T=1.55$, mass$=1$) with Lennard-Jones interaction, and periodic boundary condition in the y direction (no wall). The obstacle consists of three layers of fixed atoms with atomic separations 1 and occupying $X=[74,76], Y=[16,34]$, which interact with the fluid atoms in the same way as the fluid-fluid interaction ($R_{\rm cut}=2.5$). The reflecting probability p is set equal to 1, and to further increase the drive, *two* RPM membranes [17] are employed in series, one at $x=0$, the other at $x=3$. Over a period of 50 000 time steps, 200 000 particle data are collected, with particle positions and velocities sampled in 500 time-step intervals to eliminate redundancy due to correlation. These are then fed into the general field estimator defined in the rectangular domain (20,0)–(140,50) [18], which employs four ($=2\times2$) basis functions for the temperature field and 121 ($=11\times11$) basis functions for v_x, v_y fields. The streamlines are plotted after the continuous temperature and velocity fields have been obtained. Starting at chosen initial positions, which are equally spaced points in the up-stream region and several points in the down-stream region relative to the obstacle, the positions are determined by integrating the velocity field (symmetrized between y and $50-y$) forward in time using the leapfrog algorithm. For comparison, Fig. 10 shows the bin-averaged velocities using 60×30 mesh for the same set of particle data. The average v_x of the entire system is 0.708.

Several interesting features can be noted. The first is that the ''no-slip'' condition holds rather well near the fixed obstacle, as the velocities nearby are very small (even around the edges), and in the region occupied by the two vortices. We are confident that the vortices are not numerical artifacts because the qualitative features of these streamlines do not change when we vary the basis sets or the domain of estimation; they are also vaguely discernible on plots of meshed bin averages. Additionally, the outer streamlines that start from $x=20$ appear to be quantitatively stable against changes in the estimation procedure. What is also interesting is that the vortex configurations seem to be time dependent, i.e., they do not reach a steady state when other parts of the field do. This is reasonable in that the behavior shown in Fig. 9 certainly is not in the steady state because mass must be conserved, and there are no sources at the vortex centers. This means that using the same estimation procedure and the same number of sample data but at different time intervals during a simulation run may give different vortex shapes.

To make contact with the continuum results, the Reynolds number $\mathrm{Re}=\rho v L/\mu$, for the flow under discussion is around 10, so one may be tempted to compare with vortex generation in similar fluid mechanical scenarios [7]. However, the fact that the flow speed is very high compared to thermal velocities and that there is significant density variation along the flow means that the present situation cannot be fully treated by a continuum description.

IV. DISCUSSION

In this work two problems are addressed that pertain to molecular-dynamics simulation of fluid flow. For studying Poiseuille flow the reflecting particle method is shown to be a physically simple and computationally effective method that does not involve adding artificial body forces; its principal virtues are ease of implementation and good convergence behavior. That this method has an intrinsic mechanism

for entropy extraction is particularly noteworthy; consequently the need for temperature rescaling during simulation is eliminated. The thermodynamic field estimator is simply a systematic way to represent discrete particle data in terms of continuous distributions. Although it is presented in the present context of fluid flow simulation, the utility of this representation clearly extends to all types of discrete-particle simulations.

While each of the two methods discussed can be used separately in a variety of applications, we have in mind to combine them in formulating a general approach to couple continuum with molecular-dynamics simulation. As mentioned in the Introduction, our interest lies in a scheme, where macroscopic field boundary conditions can be imposed on an MD system in a feedback control manner. As we will present in the second paper [3] this coupling can be implemented by introducing a three-region approach. The aim is to ensure that the region of interest $\mathcal{C}$ (core) has the desired field boundary conditions *and* the natural particle dynamics. A so-called particle controller acts in an outer MD region $\mathcal{A}$, which is based on a transformation relating two distributions, one derived from the particle data (current fields) and the other from the desired fields (outer continuum solution). The core is surrounded by a buffer zone $\mathcal{B}$, which separates it from $\mathcal{A}$. The function of the field estimator developed here is to serve as the detector of the control loop, inferring the current fields at the core boundary $\partial\mathcal{C}$ based on particle data from the *entire* $\mathcal{C}$ region. A central control algorithm now compares the estimated current fields with the desired fields on $\partial\mathcal{C}$, and sends high-level instructions to the particle controller acting in $\mathcal{A}$. We will show through analytical arguments that an ''optimal'' particle controller in fact exists, in the sense that its mapping from one distribution to the other creates the least disturbance to a random variable sequence.

Once we know how to impose general field boundary conditions on a MD system in a least-disturbance or no-disturbance manner, as the above discussions outline, we can combine a continuum solver with MD simulation through the Schwarz iteration formalism [2,19]. This will enable us to study steady-state fluid flow treating the continuum and atomistic aspects on equal footing. The attempt to couple computational techniques on different scales has counterparts in other areas of simulation research [20].

ACKNOWLEDGMENTS

This work has been supported by Sandia National Laboratory. We are grateful to C.C. Wong for discussions and encouragement throughout the project. We also acknowledge helpful early discussions with N. G. Hadjiconstantinou. J.L. has received partial support from The Petroleum Research Fund, administered by the American Chemical Society.

[1] S. T. O'Connell and P. A. Thompson, Phys. Rev. E **52**, 5792 (1995).

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

[3] D. Liao, J. Li, and S. Yip (to be published).

[4] For experiments on microchannel flow see, for instance, J. Pfahler, J. Harley, H. Bau, and J. N. Zemel, in *Micromechanical Sensors, Actuators, and Systems*, edited by D. Cho *et al.* (ASME, New York, 1991), Vol. 32, p. 49; K. C. Pong, C. M. Ho, J. Q. Liu, and Y. C. Tai, in *Application of Microfabrication to Fluid Mechanics*, edited by P. R. Bandyopadhyay, K. S. Breuer, and J. C. Blechinger (ASME, New York, 1994), Vol. 197, p. 51.

[5] P. A. Thompson and M. O. Robins, Phys. Rev. A **41**, 6830 (1990).

[6] For an introduction to MD, see M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Clarendon, New York, 1987).

[7] G. K. Batchelor, *An Introduction to Fluid Mechanics* (Cambridge University Press, Cambridge, 1967).

[8] J. Koplik, J. R. Banavar, and J. F. Willemsen, Phys. Fluids A **1**, 781 (1989).

[9] M. Sun and C. Ebner, Phys. Rev. A **46**, 4813 (1992).

[10] The reduced units are length in terms of σ; energy in terms of ϵ_{ff}; mass in terms of m_f. Thus one reduced time unit is $(m_f\sigma^2/\epsilon_{ff})^{1/2}$, one reduced temperature unit is ϵ_{ff}/k_B, one reduced stress unit is ϵ_{ff}/σ^3, etc.

[11] P. A. Thompson and M. O. Robins, Science **250**, 792 (1990).

[12] *Maxwell's Demon: Entropy, Information, Computing*, edited by H. S. Leff and A. F. Rex (Princeton University Press, 1990).

[13] In a more general setup it may not be true, and perhaps we need to solve a scattering problem where $f1, f2$ is generalized to $f1(\mathbf{v}), f2(\mathbf{v})$. Here the claim holds because there is specular symmetry in the particle speed distribution and the reflection is elastic.

[14] E. S. Keeping, *Introduction to Statistical Inference* (Van Nostrand, Princeton, 1962).

[15] In reality certain fields may be known beforehand, such as $\rho(\mathbf{x})$ and $\mathbf{v}(\mathbf{x})$ in convectionless heat conduction problems.

[16] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, 2nd ed. (Cambridge University Press, Cambridge, 1992).

[17] It is not necessary to employ two RPM membranes or even $p=1$ to induce appreciable flow at this condition, but we use the example to show that it can be done. If necessary we can always have strong enough flow by placing several RPM membranes.

[18] Particle dynamics are strongly perturbed by RPM near the membrane, so it is advisable to discard those data and only estimate the fields inside the region of interest. The presence of a geometrical singularity (the barrier) causes large basis sets to be used.

[19] B. F. Smith, P. E. Bjorstad, W. D. Gropp, *Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations* (Cambridge University Press, Cambridge, 1996).

[20] Proceedings of the Workshop on Modeling of Industrial Materials: Connecting Atomistic and Continuum Scales, January 7–11, 1996, UC Santa Barbara, edited by S. Yip [J. Computer-Aided Mater. Design **3** (1996)].