Use the force! Reduced variance estimators for densities, radial distribution functions and local mobilities in molecular simulations

Benjamin Rotenberg

Use the force! Reduced variance estimators for densities, radial distribution functions and local mobilities in molecular simulations

Benjamin Rotenberg[1]

Sorbonne Universités, CNRS, Physico-Chimie des électrolytes et Nanosystèmes Interfaciaux, F-75005 Paris, France[a)]

(Dated: September 28, 2020)

Even though the computation of local properties, such as densities or radial distribution functions, remains one of the most standard goals of molecular simulation, it still largely relies on straighforward histogram-based strategies. Here we highlight recent developments of alternative approaches leading, from different perspectives, to estimators with a reduced variance compared to conventional binning. They all make use of the force acting on the particles, in addition to their position, and allow to focus on the non-trivial part of the problem in order to alleviate (or even remove in some cases) the catastrophic behaviour of histograms as the bin size decreases. The corresponding computational cost is negligible for molecular dynamics simulations, since the forces are already computed to generate the configurations, and the benefit of reduced-variance estimators is even larger when the cost of generating the latter is high, in particular with *ab initio* simulations. The force sampling approach may result in spurious residual non-zero values of the density in regions where no particles are present, but strategies are available to mitigate this artefact. We illustrate this approach on number, charge and polarization densities, radial distribution functions and local transport coefficients, discuss the connections between the various perspectives and suggest future challenges for this promising approach.

I. INTRODUCTION

Molecular Dynamics (MD) and Monte Carlo (MC) algorithms allow to sample configurations from a statistical ensemble and to numerically compute observable properties as averages over configurations for realistic models of interacting atoms or molecules, when analytical theories are usually limited to cruder models[1]. Since the early days of simulations, the tremendous development of accurate quantum and classical descriptions and of efficient sampling algorithms, supported by the ever growing availability of computational resources, established molecular simulation as an essential tool and, more fundamentally, a new scientific approach to investigate the properties of Matter[2].

The determination of the local properties in condensed matter has always been one the main applications of molecular simulation. To mention but a few examples in Biology, Chemistry and Physics, one can cite the characterization of the local structure: of water around solutes such as biological molecules by 3D densities[3–5]; of interfacial number and charge densities at electrochemical interfaces involving aqueous electrolytes or room temperature ionic liquids (RTIL)[6–9]; of the complex solvation structure in RTILs[10] or of aqueous solutes near an electrode surface[11]; of the radial distribution in concentrated electrolytes[12] in order to understand the intriguing behavior of the correlation length reported in these systems[13]. These structural properties are also used to infer thermodynamic quantities, such as local solvation entropies, energies or free energies around proteins[14], or Kirkwood-Buff integrals in mixtures[15]. Finally, the densities or radial distribution functions obtained from molecular simulations are often used as reference data to test and/or parameterize liquid state theories such as 3D-RISM[16,17] or molecular density functional theory[18–21] as well as coarse-grained models[22,23] for mesoscale simulations.

Compared to the development of models to describe the systems and of algorithms to sample statistical ensembles, surprizingly little effort has been devoted to the improvement of estimators to compute local properties such as densities or radial distribution functions from the available configurations. In practice, one generally relies on histograms couting the number of particles in a voxel of finite size h^d (in $d = 1$, 2 or 3 dimensions) around a given point or of pairs separated by a distance comprised between r and $r + \Delta r$. As discussed in more detail below, these straightforward estimators provide the correct expectation value, but behave poorly when increasing the resolution, with a diverging variance as h or $\Delta r \to 0$. The benefit of estimators with a reduced variance would be even larger when the computational cost to generate configurations is high, for example for *ab initio* MD.

Here we highlight recent developments of alternative approaches leading, from different perspectives (*e.g.* inspired from zero-variance estimators developed in Quantum Monte Carlo[24–26], or directly rooted in Statistical Mechanics[27]), to estimators with a reduced variance compared to conventional binning for densities or radial distribution functions[28–32], or even local transport coefficients[33]. A common point of all the resulting expressions is to use the force acting on the particles, in addition to their position. In Section II, we introduce the main problem of histogram-based estimators and the idea behind force sampling strategies. Section III then illustrates the variance reduction obtained from force-based estimators on local number densities in several dimensions, and extensions to charge and polarization densities, or to take into account constraints for rigid molecules. Sections IV and V are devoted to radial distribution functions and local transport coefficients, respectively. The link between these and related approaches is further discussed in Section VI, while Section VII offers some suggestions of future challenges.

[a)]Electronic mail: benjamin.rotenberg@sorbonne-universite.fr

II. FORCE SAMPLING

A. What's wrong with binning?

The computation of local densities or radial distribution functions (rdf) from molecular simulations is based on their statistical mechanical definitions as ensemble averages over microscopic configurations. Let us first consider for simplicity a system of N independent atoms, without distance or angular constraints to define rigid molecules (see section III C). A point in phase space is defined by the set of all positions $\mathbf{r}^N = \{\mathbf{r}_1, \ldots, \mathbf{r}_N\}$ and momenta $\mathbf{p}^N = \{\mathbf{p}_1, \ldots, \mathbf{p}_N\}$. The potential energy of the system, $U(\mathbf{r}^N)$, includes the interactions between atoms and the effect of external potentials. In the canonical ensemble, with fixed volume V and temperature T, the local number density in particles of type a is defined as[1]

$$\rho_a(\mathbf{r}) = \left\langle \sum_{i=1}^{N_a} \delta(\mathbf{r}_i - \mathbf{r}) \right\rangle, \qquad (1)$$

where the sum runs over atoms of type a, δ is the (3D) Dirac delta function and $\langle \ldots \rangle$ denotes an average in the canonical ensemble. Specifically, the average of an observable f is

$$\langle f \rangle = \frac{1}{\mathscr{Z}} \int d\mathbf{r}^N d\mathbf{p}^N f(\mathbf{r}^N, \mathbf{p}^N) e^{-\beta \mathscr{H}(\mathbf{r}^N, \mathbf{p}^N)}, \qquad (2)$$

where $\beta^{-1} = k_B T$ is the thermal energy, $\mathscr{H}(\mathbf{r}^N, \mathbf{p}^N) = U(\mathbf{r}^N) + K(\mathbf{p}^N)$, with $K(\mathbf{p}^N)$ the kinetic energy, is the Hamiltonian of the system, and $\mathscr{Z}$ is the partition function. Expressions similar to Eq. 1 can be written for the 1D- and 2D-densities in cartesian coordinates, using $e.g.$ $\delta(z_i - z)$ and $\delta(x_i - x)\delta(y_i - y)$. In addition, rdfs can be defined as

$$g_{ab}(r) = \frac{V}{N_a N_b} \frac{1}{4\pi r^2} \left\langle \sum_{i=1}^{N_a} \sum_{j=1}^{N_b} {}' \delta(r_{ij} - r) \right\rangle, \qquad (3)$$

where r_{ij} is the distance between particles i and j and the prime in the second sum indicates that $j = i$ should be excluded when $b = a$.

These expressions lead straightforwardly to algorithms based on a discretization of space into a grid of voxels with size h^d, with h the grid spacing and $d \in \{1, 2, 3\}$, for the density and similarly of distances, with a bin width Δr, for the rdfs. Histograms are obtained by simply counting the number of particles in a given voxel (for densities) or of particle pairs separated by a distance comprised between r and $r + \Delta r$ (for rdfs) for each configuration, summing over particles or pairs, and averaging over a collection of configurations generated according to their weight in the canonical ensemble using MD or MC simulations.

In order to illustrate the limitations of this histogram approach, which remains the most commonly used to date, let us consider a system of identical non-interacting particles in 3 dimensions. In the absence of an external potential, one expects the density to be uniform and equal to the average density $\rho = N/V$. If the bin size h is small, each voxel of the grid discretizing space will be either empty or contain

1 particle, leading to an instantaneous estimate of the density $\tilde{\rho}$ in this voxel of $0/h^3$ or $1/h^3$, respectively. In addition, the fraction of occupied voxels for each configuration, which for a sufficiently large number of configurations is also the fraction of configurations in which a given voxel is occupied, is $\alpha = \rho h^3 \ll 1$. The estimator $\tilde{\rho}$ of the local density provides the correct average over many configurations, since $\langle \tilde{\rho} \rangle = \alpha \times h^{-3} + (1 - \alpha) \times 0 = \rho$. The quality of this estimator can be measured by computing its variance: $\langle \delta \tilde{\rho}^2 \rangle = \alpha \times (h^{-3} - \rho)^2 + (1 - \alpha) \times (0 - \rho)^2 = \alpha(1 - \alpha)h^{-6} \approx \rho/h^3$. As a result, the variance diverges as the bin size h decreases.
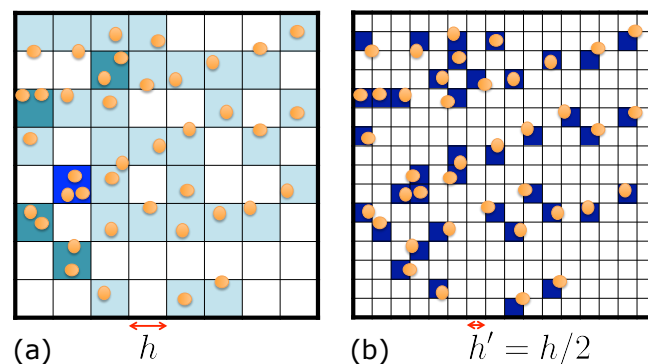

Figure 1. Contribution of a microscopic configuration, with orange disks representing atoms in a simulation box, to the number density on a grid, estimated using histograms with bins of width h (a) and $h' = h/2$ (b). The colors indicate the instantenous estimate of the local density, from the number of atoms present in each cell and the volume of the latter.

This problem of histogram-based estimates of the density is illustrated in Figure 1: When the bin size is small, most voxels are empty and the estimate of the local density is given by the fraction of configurations in which a given bin is occupied, which is small and therefore requires a large number of configurations to converge. Since the computational cost essentially comes from generating the configurations (and not from the estimate of microscopic properties from these configurations), improved estimators with a reduced variance would decrease the number of samples necessary to converge the results. The advantage would be even greater for computationally intensive approaches, based $e.g.$ on $ab\ initio$ calculations, with which data is usually scarce.

B. The idea behind force sampling

The previous discussion shows that the variance issue with histograms arises essentially from the ideal gas contribution. Fortunately, this contribution is known: It is simply the average density $\rho = N/V$. However, what really matters are the $variations$ in density with respect to the position (or in rdfs with distance), which result from the interactions between the particles and from external potentials. The idea behind force-sampling strategies is to sample these variations (the gradients) using estimators involving the force acting on the atoms

– a strategy well summarized in the title of Ref. 30: "Better than counting: density profiles from force sampling". In some cases, determining the density (or rdf) from its gradient numerically requires introducing some discretization. In others, this can be done analytically, so that the density or rdf can be determined with arbitrary resolution – in strong contrast with the conventional histogram-based methods. Before turning to the presentation and discussion of force-based estimators, we note that they fundamentally differ from a simple data-smoothing approach (*e.g.* with splines) of histogram-based ones, as they make use of additional information pertaining to the gradient of the property of interest, rather than combining values of this property computed at different positions.

III. DENSITY

We begin our survey of improved estimators using force sampling on the case the number density in Section III A, highlighting some pitfalls and how to mitigate them in Section III B, before turning to recent extensions to rigid bodies or generic densities in Section III C.

A. Variance reduction from force sampling

Starting from the definition of the density as an ensemble average, Eq. 1, and by differentiating with respect to the position $\mathbf{r}$, one obtains (omitting the species type a)

$$\nabla \rho(\mathbf{r}) = \beta \mathbf{F}(\mathbf{r}) = \beta \left\langle \sum_{i=1}^{N} \delta(\mathbf{r}_i - \mathbf{r}) \mathbf{f}_i \right\rangle, \qquad (4)$$

where we have introduced the force density $\mathbf{F}(\mathbf{r})$ expressed as an ensemble average of the force $\mathbf{f}_i$ acting on the particles, including external potentials and the interactions with other particles. The appearance of the force is due to the gradient of the Boltzmann weight, since $\nabla_{\mathbf{r}_i} e^{-\beta \mathcal{H}} = e^{-\beta \mathcal{H}} (-\beta \nabla_{\mathbf{r}_i} U) = e^{-\beta \mathcal{H}} \beta \mathbf{f}_i$ in the average Eq. 2, and the properties of the Dirac delta function. The force sampling approach hence consists in first sampling the force density from the trajectory, and then computing the density from its gradient. More precisely, this "inversion of the gradient" provides the density up to a constant, which can be taken as the average density ρ_0 to obtain the excess density $\Delta \rho(\mathbf{r}) = \rho(\mathbf{r}) - \rho_0$. Such an approach avoids the computation of the ideal gas contribution and focuses on the non-trivial contribution from interactions. In 2 or 3 dimensions, the gradient can formally inverted as[30]

$$\nabla^{-1} \cdot \mathbf{F}(\mathbf{r}) = \frac{1}{c_d} \int d\mathbf{r}' \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^d} \cdot \mathbf{F}(\mathbf{r}') \qquad (5)$$

with $c_d = 4\pi$ if $d = 3$ and $c_d = 2\pi$ if $d = 2$. In practice, this integration of the gradient can also be done efficiently in reciprocal space, using Fast Fourier Transforms (FFT), from the force density discretized on a grid (see Refs 29,32).
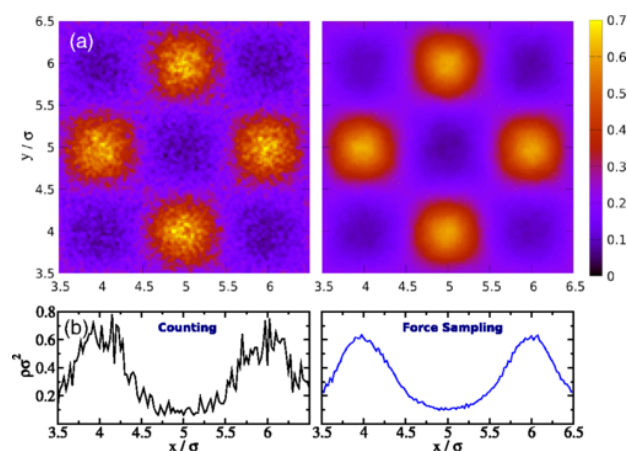

Figure 2. (a) 2D density for Lennard-Jones particles in an external potential $V_{ext}(\mathbf{r}) = V_0 \sin(2\pi n_w x/L) \sin(2\pi n_w y/L)$ with $V_0 = \varepsilon$, the LJ energy, $n_w = 5$ and $L = 10\sigma$, with σ the LJ diameter. Simulations are performed with 25 particles and only the central region of the box is shown. Results are obtained from 10^6 configurations using histograms (left) and force sampling (right); bins are squares of side length 0.025σ. (b) Density profile as a function of x at constant $y = 5\sigma$ from histograms (left) and force sampling (right). Reprinted with permission from de Las Heras and Schmidt, *Phys. Rev. Lett.* 2018, **120**, 218001. Copyright (2020) by the American Physical Society.

Figure 2 illustrates the benefit of the force sampling approach on a 2D system of Lennard-Jones particles in an external potential, from de Las Heras and Schmidt[30]. Both panels, showing respectively a 2D density map and a 1D cut through the latter, clearly demonstrate that using data from the same configurations from MC simulations and the same grid, the force-based estimator displays much less noise than the direct histogram approach. As expected from the discussion of Section II, the error with respect to the converged result with many configurations diverges as the bin size decreases and behaves much better with force scaling (see Ref. 30 for a discussion of the scaling on a 1D example, and Ref. 32 on a 3D example). While force sampling provides a clear reduction of variance compared to histograms for small bins, the benefit is less obvious for larger ones (a more precise statement would have to be system specific), in particular because the numerical integration of the gradient on a coarse grid also introduces some discretization errors.

B. Pitfalls and mitigation

Another point deserving the attention of the reader is that the force sampling approach may result in non-zero values of the density in regions where no particles are present. This is illustrated on a 1D example in Figure 3, from Purohit *et al.*[31], which compares the conventional binning approach to force sampling as described above and to mapped averaging, which is a more general framework that can be applied to recover force sampling methods for density distributions (see below). All three methods provide similar results everywhere,
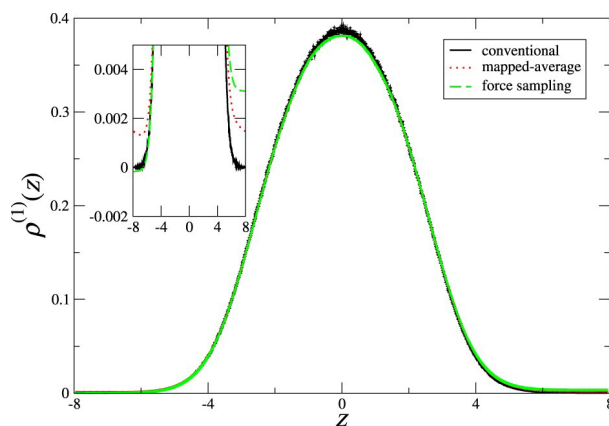
Figure 3. Density profile for a Lennard-Jones fluid in an external potential $\varepsilon(z/\sigma)^2$, with σ and ε the LJ diameter and energy. The results obtained from Monte Carlo simulations are shown in units $\sigma = 1$ with three estimators: conventional histograms, force sampling and mapped-averaging (see Eq. 6 for the latter). The inset shows the same data on a scale allowing a better visualization of the tails of the distributions. Reproduced from Purohit *et al.*; *Molecular Physics* 2019, **117**, 2822, with permission of Taylor & Francis Ltd.

but the inset shows that they differ in the tails of the distribution. While with histograms the density falls to zero when the external potential diverges, this is not exactly the case for methods based on the force (which, on the positive side, display a lower variance): by integrating the density gradient as described above, the density starts from 0 on the left but integrates to a finite plateau on the right; mapped averaging provides a symmetrized but non-vanishing result. Fortunately, this spurious residual density, which comes from inaccuracies in the force density, can be reduced simply by increasing the number of configurations used to sample the latter.

Despite the above limitation (which also applies in higher dimensions), the 1D case also offers an example where the gradient can be integrated analytically. The above-mentioned mapped averaging framework provides a way to introduce approximate theoretical results (such as an *a priori* estimate of a density) and obtain an exact expression of the error in the theory[34,35]. In the present 1D geometry, introducing a uniform density as a prior guess, Purohit *et al.* obtained the following expression[31]

$$\rho(z) = \frac{N}{V} - \left\langle \frac{1}{S} \sum_{i=1}^{N} \left(\frac{1}{2} - H(z - z_i) - \frac{z_i - z}{L} \right) \beta f_{z,i}, \right\rangle \quad (6)$$

with L the length of the system in the z direction and $S = V/L$ the area of system in the lateral directions, H the Heaviside function, and $f_{z,i}$ the z-component of the force acting on particle i. It is probably also possible, using a different prior estimate, to recover from this formalism another expression obtained in Ref. 33 from a different perspective, namely to combine a position-based and a force-based estimator. To that end, one can introduce appropriate weights $w_N(z)$ and $w_f(z)$, where the subscripts refer to number and force, respectively, such that $w_f'(z) = \delta(z) - w_N(z)$ and that $w_f(z)$ vanishes when

$|z|$ is large, in order to write the 1D number density as:

$$\rho(z) = \frac{1}{S} \left\langle \sum_{i=1}^{N} \delta(z_i - z) \right\rangle$$

$$= \frac{1}{S} \left\langle \sum_{i=1}^{N} w_N(z_i - z) \right\rangle - \frac{\beta}{S} \left\langle \sum_{i=1}^{N} f_{z,i} w_f(z_i - z) \right\rangle \quad (7)$$

The function $w_N(z)$ can be seen as a coarse-graining kernel for the contribution of each particle to the number density and should therefore vanish beyond a coarse-graining length ξ. Families of estimators can be obtained by choosing various forms for $w_N(z)$ and corresponding $w_f(z)$, and varying ξ (see Ref. 33 for an example of weight functions and a discussion of the choice of ξ). This provides a handle to mitigate the artefact of non-zero density in empty regions predicted by integrating the force density. The possible connection with mapped averaging can be hypothesized from the observation that Eq. 7 reduces to Eq. 6 (which is a particular case of a more general expression[31]) for $w_N(z) = 1/L$ (which strictly speaking does not satisfy the criteria to be interpreted as a coarse-graining kernel, since every particle would contribute to the density everywhere). In practice, the estimator defined in Eq. 7 can be computed efficiently by convoluting *a posteriori* the histogram-based estimators of the number and force densities with their corresponding weight functions.

C. Extensions: rigid bodies, generic densities

The force sampling approach to number density has recently been extended in several directions by Coles *et al.*[32]. From the practical point of view, it is important to consider the case of molecules described as rigid bodies using distance constraints, which includes popular water models such as SPC/E or the TIPnP family. Even though the derivation is not straightforward, the final result is particularly simple[32]:

$$\nabla \rho(\mathbf{r}) = \beta \mathbf{F}(\mathbf{r}) = \beta \left\langle \sum_{i=1}^{N} \delta(\mathbf{r}_i - \mathbf{r}) \mathbf{f}_i^* \right\rangle_{constr.}, \quad (8)$$

where $\mathbf{f}_i^*$ is *the sum of forces acting on particle i and all particles participating in a constraint with i* (in particular those belonging to the same rigid molecule as i) and where the average is made over configuration satisfying all constraints. In practice, the gradient of the water oxygen (resp. hydrogen) density is computed by assigning the total force acting on each molecule to the position of the O atom (resp. H atoms).

Another generalization was to consider other local quantities such as the charge or polarization densities. Such an extension is trivial for combinations of the type

$$A(\mathbf{r}) = \left\langle \sum_{i=1}^{N} \delta(\mathbf{r}_i - \mathbf{r}) a_i \right\rangle_{constr.} \quad (9)$$

when the microscopic property a_i does *not* depend on the coordinates $\mathbf{r}^N$, which is the case for the charge density. Following the same derivation as for the number density, one readily

obtains:

$$\nabla A(\mathbf{r}) = \beta \left\langle \sum_{i=1}^{N} \delta(\mathbf{r}_i - \mathbf{r}) a_i \mathbf{f}_i^* \right\rangle_{constr.} \tag{10}$$

Unlike the charge density, the polarization density depends on the orientation of the molecules, hence their atomic coordinates. For rigid polar molecules, it is nevertheless possible to use a different set of coordinates, namely the positions of their centers of mass $\mathbf{R}^{N_r}$ and orientations $\mathbf{\Omega}^{N_r}$, with N_r the number of rigid molecules. The result for the α component of the polarization density $\mathbf{P}(\mathbf{r})$ is:

$$\nabla P_\alpha(\mathbf{r}) = \beta \left\langle \sum_{l=1}^{N_r} \delta(\mathbf{R}_l - \mathbf{r}) \mu_{l,\alpha} \mathbf{f}_l^* \right\rangle \tag{11}$$

where the sum runs over molecules l and $\mu_{l,\alpha}$ is the α component of their molecular dipole. In that case there is no need to include constraints explicitly in the average since there are no additional degrees of freedom.
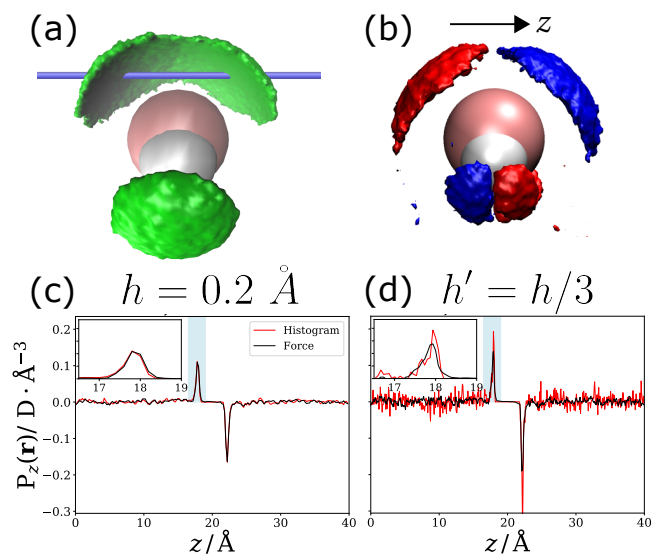

Figure 4. Water structure around a fixed water molecule, from simulations with the rigid SPC/E water model. (a) Number density; the isosurface bounds the regions where the density is greater than 0.07 Å^{-3} and illustrates the position of water molecules in the first solvation shell. The blue line indicates the position of the axis along which the data of panels (c) and (d) are shown. (b) z-component of the polarization density; blue and red isosurfaces bound regions where the z-component of the polarization density is less than -0.035 D Å^{-3} and greater than $+0.035$ D Å^{-3}, respectively. The remaining panels show the z-component of the polarization density for the single line of voxels illustrated in panel (a), using histograms (red) or force sampling (black), using a grid spacing $h = 0.2$ Å (c) or $h' = h/3$ (d). The insets are close-ups on the rising edge highlighted in blue on the main figures. Adapted from *J. Chem. Phys.* 2019, **151**, 064124, with the permission of AIP Publishing.

Figure 4 illustrates results on the 3D organization of water, around a fixed water molecule from simulations with the rigid SPC/E water model, with results from Ref. 32. Panel 4a shows

the regions of large number density of O atoms, computed using Eq. 8 on a grid and integrating the gradient numerically using FFT. One can identify the position of the first solvation shell, with different basins corresponding to molecules donating (top) or receiving (bottom, only one is visible from this angle) H-bonds to/from the central molecule. Panel 4b then shows the z component of the polarization density, computed using Eq. 11 and FFT to integrate the gradient. This allows clarifiying the orientation of the molecular dipole of water molecules in the different basins of panel 4a, consistent with the donation/reception of H-bonds, and shows in particular that P_z vanishes in the plane of the central molecule, as expected from the symmetry of the system.

In order to demonstrate the benefit of force sampling compared to histograms to determine the 3D number and polarization densities, panels 4c and 4d finally show a 1D trace through the 3D polarization density along the blue line shown in panel 4a. Both methods display two peaks of opposite sign in the vicinity of the molecule, corresponding to the two basins for H-bond donating molecules of panel 4b. For a relatively large bin size, both methods give comparable results. As the bin size decreases, the quality of the histogram-based estimator deteriorates much faster than the force-based estimator: This is visible in the larger amplitude of the noise far from the solute, as well as near the rising edge close to the latter (see the insets) and the growing asymmetry between the two sides of the molecule.
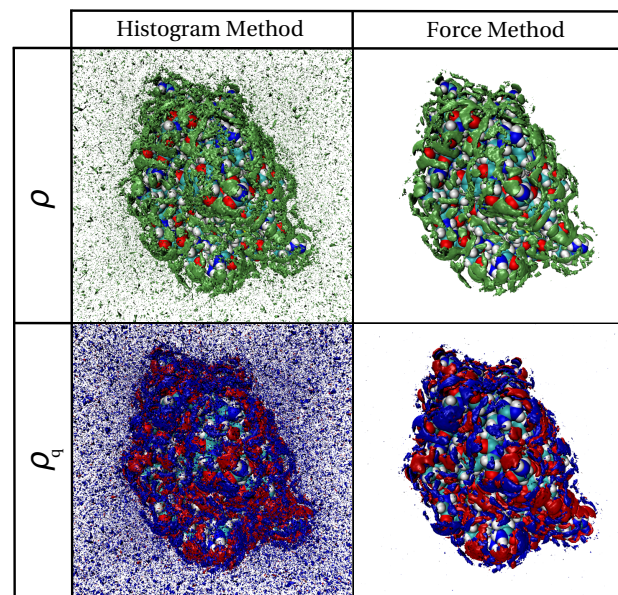

Figure 5. Isosurfaces for the number density ρ (top) and the charge density ρ_q (bottom) around a lysozyme protein in water, as obtained from histograms (left) or force sampling (right). The green surfaces bound the region where the number density is greater than 0.1 Å^{-3}, while the blue (negative) and red (positive) surfaces bound regions where the magnitude of the charge density exceeds ± 0.1 e Å^{-3}. Reprinted from *J. Chem. Phys.* 2019, **151**, 064124, with the permission of AIP Publishing.

As a final illustration on a more complex system, Figure 5,

also from Ref. 32, compares the number and charge densities around a small protein, lysozyme. Following Eq. 10, the position and charge q_i of each atom is considered when computing the gradient, but the relevant force is the total force acting on the rigid molecule. The advantage of force sampling over histograms is again clearly visible in the reduction of the noise away from the solute, as well as in a better resolved solvation structure near its surface. Among other features, the positive and negative lobes of the charge density allow one to identify molecules donating or receiving H-bonds.

IV. RADIAL DISTRIBUTION FUNCTION

While the previous section only considered one-body densities, other statistical tools exist to quantify the structure of a system. We now turn to the force sampling approach to radial distribution functions, which reflect correlations between particles.

A. Virial-like estimator

Inspired by ideas for the electron density in Quantum Monte Carlo[24–26], Borgis *et al.*[29] proposed an expression alternative to the definition Eq. 3. Starting from the latter, they used Poisson's equation to rewrite $\delta(r) = -\frac{1}{4\pi}\Delta\frac{1}{r}$, with Δ the Laplace operator. Integration by parts in the canonical average, symmetrization beteween particles i and j, and finally Gauss's law then lead to:

$$g_{ab}(r) = 1 + \frac{\beta}{4\pi}\frac{V}{N_aN_b}\left\langle \sum_{i=1}^{N_a}\sum_{j=1}^{N_b}{}' \frac{1}{2}(\mathbf{f}_i - \mathbf{f}_j)\cdot\frac{\mathbf{r}_j - \mathbf{r}_i}{r_{ij}^3}H(r_{ij} - r)\right\rangle$$

$$(12)$$

with H the Heaviside function. The origin of the forces $\mathbf{f}_i$ and $\mathbf{f}_j$ acting on the particles is again the gradient of the Boltzmann weight with respect to the particles positions, which comes from integrating by parts. This expression displays naturally the separation between the ideal gas contribution, which as described in Section II A is detrimental to the convergence of the histogram-based estimator, and a virial-like correction arising from interactions between particles and external potentials. It is worth emphasizing here that the above expression holds even with many-body potentials, since no assumption of pairwise additivity was needed in the derivation.

One can further note that it involves a Heaviside instead of a Dirac delta in Eq. 3. This results from the integration by parts in the canonical average, which can be seen as a way to "integrate the gradient" of g analytically, and has important practical consequences. While in the histogram approach each pair only contributes to the estimate of the rdf in the bin corresponding to $r = r_{ij}$ and symmetrically the estimate at r only benefits from pairs such that $r_{ij} = r$, with Eq. 12 each pair contributes to the estimate of the rdf for all values $r \le r_{ij}$ and symmetrically the estimate of $g(r)$ benefits from all pairs separated by a distance larger than r. In addition, bins are not

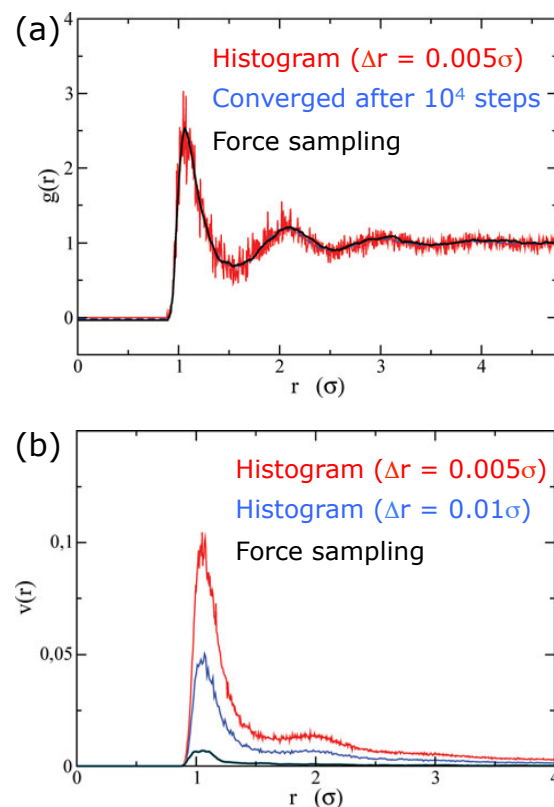necessary anymore and one can comptute the rdf with arbitrary resolution in r.

Figure 6. Radial distribution function (rdf) for a Lennard-Jones (LJ) fluid at reduced density $\rho^* = 0.8$ and reduced density $T^* = 1.35$, from a simulation with 864 particles. (a) rdf obtained from a single microscopic configuration with histograms (red) with a small bin width $\Delta r = 0.005\sigma$, with σ the LJ diameter, and with force sampling (black); the latter is indistinguishable from the converged result with histograms averaged over 10^4 configurations (blue). (b) Variance of the estimators, for each distance r, over 10^3 configurations: histograms with $\Delta r = 0.005\sigma$ (red) and $\Delta r = 0.01\sigma$ (blue), and force sampling (black). Reproduced from Borgis *et al.*, *Molecular Physics* 2013, **111**, 3486, with permission of Taylor & Francis Ltd.

The benefit of the force sampling approach to estimate rdfs is illustrated in Figure 6, which reproduces results from Ref. 29 for a Lennard-Jones fluid. Panel 6a shows that the estimate obtained on a single configuration with histograms (using a very small bin width) is much noisier than with force sampling, the latter being indistinguishable from the converged histogram-based estimate from 10^4 configurations. Panel 6b then shows the variance of the estimator over $N_{conf} = 10^3$ configurations as a function of r, defined by

$$v(r) = \frac{1}{N_{conf}}\sum_{k=1}^{N_{conf}}g_k(r)^2 - \left(\frac{1}{N_{conf}}\sum_{k=1}^{N_{conf}}g_k(r)\right)^2 \quad (13)$$

The shape of $v(r)$ is the same for all methods, in particular it vanishes inside the core (where g should vanish) and

at long distance (where $g \approx 1$). However, the variance of the force-based estimator is significantly reduced compared to histograms, even more so that the bin size is small for the latter.

B. Alternative derivations and expressions

As mentioned earlier, Eq. 12 can be seen as an integration of $g'(r) = dg/dr$ from $r \to \infty$ where $g = 1$, leading to two difficulties. Firstly, in a finite box the limiting value of g is never reached exactly. In fact for simulations in the canonical ensemble, the plateau value differs from 1 by a $\mathcal{O}(N^{-1})$ correction[36,37]. Secondly, this expression does not guarantee *a priori* that g will vanish inside the core, which is the case for $r = 0$ only if

$$\frac{\beta}{4\pi}\frac{V}{N_a N_b}\left\langle \sum_{i=1}^{N_a}\sum_{j=1}^{N_b}{}' \frac{1}{2}(\mathbf{f}_i - \mathbf{f}_j)\cdot\frac{\mathbf{r}_j - \mathbf{r}_i}{r_{ij}^3}\right\rangle = -1 . \qquad (14)$$

In practice, this is almost the case (hence it is hard to see the small but finite value in Figure 6a), but only approximately. For systems with sufficiently hard repulsion at short range, one can enforce that $g(0) = 0$ and integrate from $r = 0$ to obtain an alternative expression

$$g_{ab}(r) = \frac{\beta}{4\pi}\frac{V}{N_a N_b}\left\langle \sum_{i=1}^{N_a}\sum_{j=1}^{N_b}{}' \frac{1}{2}(\mathbf{f}_i - \mathbf{f}_j)\cdot\frac{\mathbf{r}_j - \mathbf{r}_i}{r_{ij}^3}H(r - r_{ij})\right\rangle . \qquad (15)$$

A similar expression was obtained for the two-body density by Purohit *et al.* using mapped averaging (already discussed for the one-body density in Section III)[31]. Compared to Eq. 12, each pair now contributes to the estimate of the rdf for all values $r \geq r_{ij}$ and symmetrically the estimate of $g(r)$ benefits from all pairs separated by a distance smaller than r.

It is worth noting that another similar expression had in fact been introduced earlier by Adib and Jarzynski[28] as an unbiased estimator first for the density around a fixed solute, then for the rdf. Starting from the expression of a local (surface) quantity in terms of a volume average, using Gauss's theorem and introducing an appropriate vector field $\mathbf{u}$, they obtained an expression of the rdf (Eq. 20 of Ref. 28) valid for pairwise additive potentials – while the above expressions do not rely on such an approximation and can be used also with many-body potentials such as in *ab initio* simulations. The vector field $\mathbf{u}$ introduced in this work is similar to the $\nabla\frac{1}{r} = -\frac{\mathbf{r}}{r^3}$ in the above expressions, but also accounts for the possible presence of a hard sphere solute at the origin and introduces the largest sphere that fits in the simulation box ($R_{max} = L/2$ with L the box length). It might therefore be useful to consider a similar quantity in the derivation of Ref. 29.

As a final remark on rdfs, we also mention that the idea of integrating by parts in the canonical average can be pushed even further. In the same study[29], Borgis *et al.* obtained by performing a second integration by parts another expression

of the rdf:

$$g_{ab}(r) = 1 + \frac{\beta}{4\pi}\frac{V}{N_a N_b}\left\langle \sum_{i=1}^{N_a}\sum_{j=1}^{N_b}{}' \frac{1}{2}(\Phi_i + \Phi_j)\min\left(\frac{1}{r},\frac{1}{r_{ij}}\right)\right\rangle \qquad (16)$$

with $\Phi_i = \beta\mathbf{f}_i^2 - \Delta_{\mathbf{r}_i}U$. Unlike Eq. 12 in which only the force enters, this one requires the Laplacian of the energy (trace of the force gradient) with respect to $\mathbf{r}_i$, which is usually not computed in molecular dynamics simulations. In addition, preliminary numerical tests in this study proved disappointing. However, this illustrates the fact that there are many possibilities to obtain alternative estimators (even though not always practically useful, at least immediately).

V. LOCAL TRANSPORT COEFFICIENTS

Beyond structural properties, the force sampling strategy has also been explored recently for the computation of local transport coefficients in confined fluids[33], such as the system illustrated in Figure 7a. Due to the external potential from the walls, the fluid adopts a layered structure, shown in Figure 7b. In the presence of an external perturbation parallel to the walls, such as a pressure gradient $-\nabla P$ or a chemical potential gradient $-\nabla\mu$ (or an electric field $-\nabla\psi$, *e.g.* for an electrolyte solution) the various components of the fluid will respond differently and the local steady-state flux of each species will depend on the type of perturbation. For sufficiently small perturbations, the response is linear and fully characterized by a *mobility matrix* $\mathcal{M}(z)$ relating local fluxes to the external forces. For the binary fluid of Figure 7, this can be written as

$$\begin{pmatrix} q(z) \\ j_A(z) - c_A^* q(z) \end{pmatrix} = \mathcal{M}(z)\begin{pmatrix} -\nabla P \\ -\nabla\mu \end{pmatrix} , \qquad (17)$$

where q and j_A are the local volume fluxes and solute fluxes, respectively, and c_A^* is a reference concentration related to the composition in the bulk part of the fluid. For example, the element $\mathcal{M}_{11}(z)$ of this 2×2 matrix predicts the flux of all species under a pressure gradient (Poiseuille flow), while $\mathcal{M}_{12}(z)$ corresponds to diffusio-osmosis.

A. Mobility profiles from equilibrium molecular dynamics

For illustration purposes, we will only consider here $\mathcal{M}_{11}(z)$, but the other cases can be found in Ref. 33. Using linear response theory, it is possible to derive expressions of the mobility profiles $\mathcal{M}_{ij}(z)$ as integrals of correlation functions (Green-Kubo approach, GK)

$$\mathcal{M}_{11}^{GK}(z) = \frac{V}{k_B T}\int_0^{+\infty} dt\, C_{qQ}(t,z) , \qquad (18)$$

with $C_{qQ}(t,z) = \langle q(z,t)Q(0)\rangle$ is a time-correlation function (for each z) between the local volume flux

$$q(z,t) = \frac{H}{N}\sum_{i=1}^{N} v_{x,i}(t)\delta(z_i(t) - z) , \qquad (19)$$
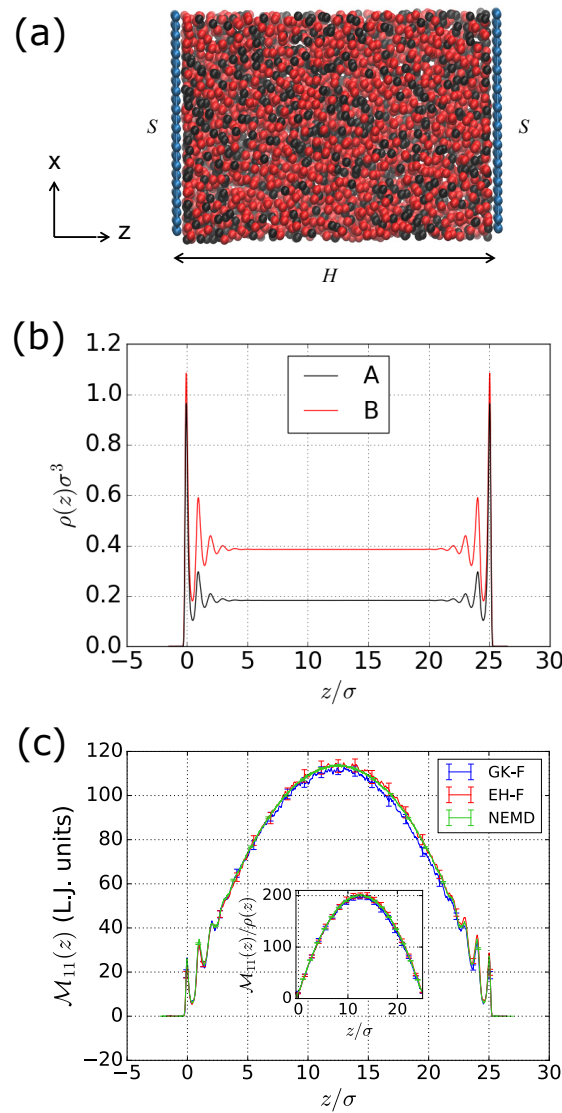
Figure 7. (a) Binary Lennard-Jones (LJ) fluid confined between walls consisting of LJ particles. The solute A (black) and solvent B (red) are identical but the solute has a stronger interaction with the wall. (b) Density profiles; the position and density are given in units σ and σ^{-3}, respectively, with σ the LJ diameter common to all species. (c) Mobility profile $\mathscr{M}_{11}(z)$ describing the linear response of the total flux to a pressure gradient (see Eq. 17), obtained from equilibrium simulations using force sampling with Green-Kubo (GK) and Einstein-Helfand (EH), as well as from non-equilibrium molecular dynamics (NEMD). The inset shows that $\mathscr{M}_{11}(z)/\rho(z)$ exhibits the parabolic shape expected from continuum hydrodynamics for the velocity profile in this case (Poiseuille flow). Adapted from *J. Chem. Phys.* 2020, **153**, 044125 , with the permission of AIP Publishing.

with $v_{x,i}$ the x-component of the velocity of particle i, $N = N_A + N_B$ the total number of particles and H the distance between the wall, and the global flux averaged over the fluid slab

$$Q(t) = \frac{1}{H}\int_0^H \mathrm{d}z\, q(z,t) = \frac{1}{N}\sum_{i=1}^N v_{x,i}(t) . \quad (20)$$

Alternatively, the integrals can be computed as (Einstein-Helfand approach, EH)

$$\mathscr{M}_{11}^{EH}(z) = \frac{V}{k_B T}\lim_{t\to+\infty}\frac{\left\langle \int_0^t \mathrm{d}t''\, q(z,t'') \int_0^t \mathrm{d}t'\, Q(t')\right\rangle}{2t} . \quad (21)$$

B. Force sampling for time correlation functions

Since the time correlations are defined as canonical averages of observables involving a Dirac delta, one can follow the ideas developed for static properties. From the definition

$$C_{qQ}(t,z) = \left\langle Q(0)\frac{H}{N}\sum_{i=1}^N v_{x,i}(t)\delta(z_i(t)-z)\right\rangle , \quad (22)$$

one can introduce a force-weighted observable,

$$F_{qQ}(t,z) = \left\langle Q(0)\frac{H}{N}\sum_{i=1}^N v_{x,i}(t)f_{z,i}(t)\delta(z_i(t)-z)\right\rangle , \quad (23)$$

and form, in the spirit of Eq. 7, the mixed estimator

$$\tilde{C}_{qQ}(t,z) = \int_0^H \mathrm{d}z'\left[w_N(z'-z)C_{qQ}(t,z')-w_f(z'-z)\beta F_{qQ}(t,z')\right]$$
$$= (w_N * C_{qQ})(t,z)+\beta(w_f * F_{qQ})(t,z) , \quad (24)$$

where the convolution products are in space only, not time. A similar construction can be followed for the EH approach.

Figure 7c compares the predictions from equilibrium MD simulations with the GK and EH approaches, both using the mixed estimators introduced above, for the $M_{11}(z)$ element of the mobility matrix, corresponding to a Poiseuille flow. The results are found to be in excellent agreement with the more direct non-equilibrium MD approach. The inset further shows that when considering the velocity profile, *i.e.* the flux divided by the density, one recovers the parabolic shape expected from continuum hydrodynamics. The fact that the NEMD results are recovered validates the relevance of the approach. This is particularly important because the derivation leading to the exact result for the density is complicated for time-correlation functions by the fact that observables are considered at two times: The canonical average corresponds to points in phase space at the initial time 0, where the global flux is also considered, while the local fluxes are computed from the positions and velocities at subsequent time t (it is possible to write a symmetric expression). The integration by parts over the initial positions $z_i(0)$ then introduces an additional term, $\left\langle Q(0)\frac{H}{N}\sum_{i=1}^N \frac{\partial v_{x,i}(t)}{\partial z_i(0)}w_f(z_i-z)\right\rangle$, which involves the derivative of the x component of the velocity at time t with respect to the initial position in the z direction. Although this term might not vanish in principle, it would be very difficult to evaluate from the trajectories. It was found numerically that it was sufficient to neglect it, but a formal derivation would be desirable.

The improvement of the forced-based estimator compared to histograms is discussed in detail in Ref. 33, together with results for the other elements of the mobility matrix. We

only emphasize here that such an improvement is essential, because determining the mobility profiles is computationally much more demanding than static properties: It requires converging, *for all positions*, the time correlation functions with sufficient accuracy to compute the integral in Eq. 18 (GK) or the slope in Eq. 21 (EH). To conclude this discussion of local transport properties, we note that for the determination of a single mobility profile NEMD remains more efficient than the above equilibrium MD route. However, the latter can provide all elements of the mobility matrix simultaneously, whereas NEMD requires separate simulations with the different perturbations, and to consider different magnitudes in order to check the validity of the linear response. Therefore, one can expect that the equilibrium approach will be particularly helpful for multicomponent systems, *e.g.* for the response of electrolyte solutions to pressure, chemical potential or electric potential gradients (hence a 3×3 mobility matrix). The force based estimator, with a reduced variance compared to histograms, will contribute to decreasing the number and length of trajectories necessary to obtain reliable results.

VI. DISCUSSION

In the previous sections, we have mentioned several approaches arriving from different perspectives to similar expressions for estimators with a reduced variance compared to histograms, which involve the forces acting on the particles. The examples making use of the Poisson equation to rewrite the Dirac delta and/or using Gauss's theorem involve integration by parts in the definitions as canonical averages, very much in the spirit of the Yvon theorem for averages of the derivative of an arbitrary function of the particle coordinates $A(\mathbf{r}^N)$ with respect to the position z_i of one particle[1]:

$$\left\langle \frac{\partial A(\mathbf{r}^N)}{\partial z_i} \right\rangle = \beta \left\langle A(\mathbf{r}^N) \frac{\partial U(\mathbf{r}^N)}{\partial z_i} \right\rangle = -\beta \left\langle A(\mathbf{r}^N) f_{z,i} \right\rangle .$$
(25)

They are also directly related to the notion of potential of mean force (PMF), as discussed for the rdf in Ref. 29 or the derivation of Ref. 32 for the density gradient in the presence of constraints, which makes use of the fundamental result of Ciccotti *et al.* for the mean force associated with generic collective variables[38], by choosing the 3 spatial coordinates as collective variables. The same strategy could be used to extend Eq. 12 in the presence of constraints, using the distance between particles as collective variable. One should keep in mind, however, that the force density $\mathbf{F}(\mathbf{r}) = k_B T \nabla \rho(\mathbf{r})$ (see Eq. 4) and the mean force, derived from the PMF, differ by a factor $\rho(\mathbf{r})$. A more detailed discussion of this difference and the practical consequences for their computation in molecular simulations, in the case of the rdf, can be found in Ref. 29.

The separation between the ideal gas reference and the contribution from interactions and external potentials is natural in the framework of mapped averaging, which provides the exact correction to a theory (*e.g.* uniform density). This idea of using a reference distribution, leading to estimators involving forces, had also been proposed by Basner and Jarzynski

to compute binless PMFs, with an illustration on an angular distribution[39], or Zhang and Ma to compute rdfs or angular distributions[40]. One could envision for example to apply it also in the case of the local transport properties illustrated in Section V, using the theoretical prediction from continuum hydrodynamics as reference.

Since in molecular dynamics simulations the forces are computed to generate the trajectory, the force-based estimator does not entail a larger computational cost compared to histograms. As a result, one can only recommend the use force-based estimators to compute local densities or radial distribution functions. For Monte Carlo, the additional cost of computing the forces in addition to the energy should be compared to that of generating enough configurations to achieve the same variance reduction with histograms, which depends in particular on the bin width. Another issue related to the computation of forces is the case of particles described by hard cores. Indeed, the expressions reported here cannot be used directly. However, the effect of excluded volume can be treated separately from other interactions, which allows to apply the ideas of force sampling even in this case[28,35].

VII. SUMMARY AND FUTURE CHALLENGES

Even though the determination of local properties, such as densities or radial distribution functions, remains one of the most standard goals of molecular simulation, it seems that the community still essentially relies on histogram-based algorithms to estimate them. The main objective of the present work was to highlight recent developments of alternative approaches leading, from different perspectives, to estimators with a reduced variance compared to conventional binning. They all make use of the force acting on the particles, in addition to their position, and allow to focus on the non-trivial part of the problem in order to alleviate (or even remove in some cases) the catastrophic behaviour of histograms as the bin size decreases. The corresponding computational cost is negligible for molecular dynamics simulations, since the forces are already computed to generate the configurations, and the benefit of reduced-variance estimators will be even larger when the cost of generating the latter is high, in particular with *ab initio* simulations. The force sampling approach may result in spurious residual non-zero values of the density in regions where no particles are present, but strategies are available to mitigate this artefact.

Up to now, the variance reduction via force sampling has mainly been demonstrated numerically – even though the methods themselves were of course derived as rigorously as possible and good reasons were proposed to explain the observations. Since variance reduction is a field of research by itself, computational physicists, chemists and biologists would certainly benefit from insights from the Mathematics community in order to obtain analytical predictions of the gain to be expected with respect to binning, in particular the scaling with the bin size (when the force density is first sampled on a grid, then integrated numerically) and the number of configurations. Other directions could include improved or optimal

methods to "integrate the gradient", in order to reconstruct the density from the force density, in particular in 3D, or the formal derivations in the case of transport (see the discussion of the "missing term" in Section V), including in the presence of constraints for rigid bodies[41].

The potential of these force sampling strategies could also be investigated for other thermodynamic conditions, *e.g.* the isothermal-isobaric (NPT) or the grand-canonical (μVT) ensembles, or for quantities derived from the ones considered here, *e.g.* to improve the convergence of Kirkwood-Buff integrals, which involve rdfs[42,43]. Other local properties might also benefit from the ideas behind force sampling, in particular the local stress tensor[44,45]. However, one should keep in mind that this might require quantities that are not computed in typical simulations, such as the force gradient, and that the additional computational cost should not exceed that of reducing the variance by the same amount using histograms simply by generating more configurations. Another possible direction for the short-term development of force sampling, in particular following the approach of Ref. 32 based on the results of Ref. 38, is the extension to non-cartesian coordinates and/or multidimensional collective variables, which can be used to characterize the 3D structure of liquids (see *e.g.* Ref. 46 for a recent example).

Finally, it would be interesting to explore the connections with other simulation contexts where the strategies discussed in the present work could also provide alternative expressions for observables of interest. For example, improved estimators involving the force have already been proposed in path integral molecular dynamics to estimate thermodynamic quantities, based on a virial form (see *e.g.* Refs. 47,48). In the other direction, the ability to reduce errors in the computation of densities/PMFs would also be useful for coarse-graining strategies based *e.g.* on Boltzmann inversion[22] or on the relative entropy[23], or for simulations with adaptive resolution introducing a thermodynamic force density in the hybrid region[49]. As for dynamic properties, beyond the equilibrium route presented in Section V, it could be useful to extend the force-sampling approach to non-equilibrium steady state, as suggested in Ref. 30, or even to path sampling methods[50]. We hope that the present perspective will motivate others to embark on this promising path.

DATA AVAILABILITY

Data sharing is not applicable to this article as no new data were created or analyzed in this study.
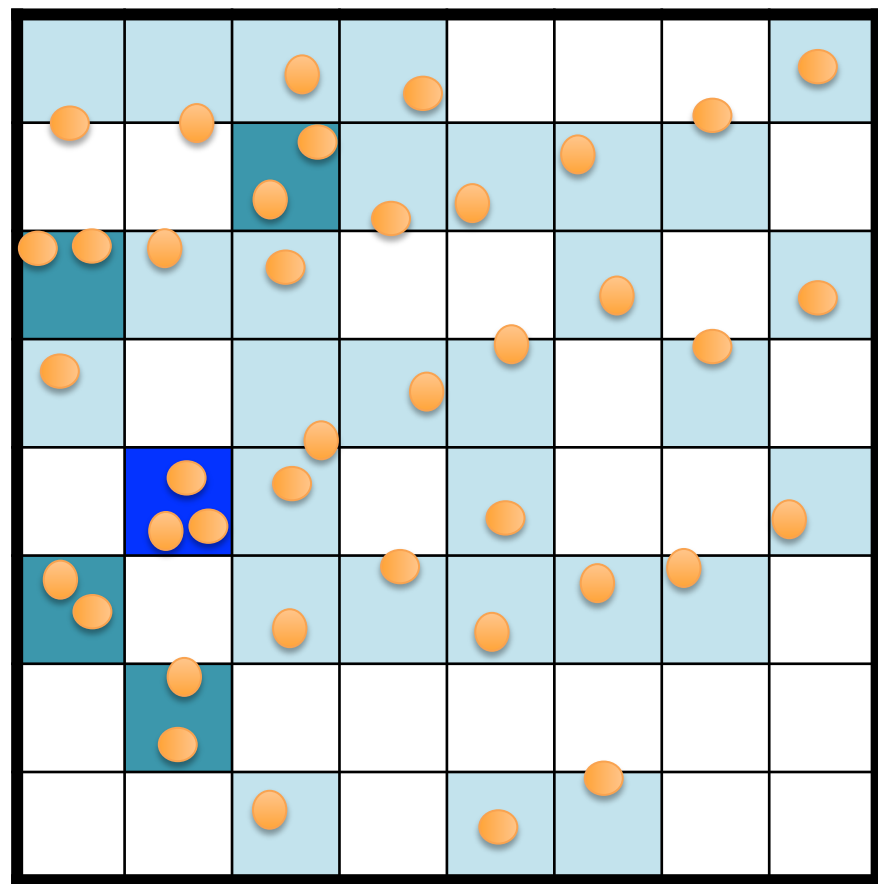
REFERENCES

[1]J. P. Hansen and I. R. McDonald, *Theory of Simple Liquids*, 4th ed. (Elsevier, Amsterdam, 2013).
[2]G. Battimelli, G. Ciccotti, and P. Greco, *Computer Meets Theoretical Physics, The New Frontier of Molecular Simulation*, 1st ed. (Springer Nature, Switzerland, 2020).
[3]R. Abel, R. A. Friesner, T. Young, B. Kim, and B. J. Berne, "Motifs for molecular recognition exploiting hydrophobic enclosure in protein–ligand binding," Proc. Natl. Acad. Sci. **104**, 808–813 (2007).
[4]I. Altan, D. Fusco, P. V. Afonine, and P. Charbonneau, "Learning about Biomolecular Solvation from Water in Protein Crystals," J. Phys. Chem. B **122**, 2475–2486 (2018).
[5]M. E. Wall, G. Calabró, C. I. Bayly, D. L. Mobley, and G. L. Warren, "Biomolecular Solvation Structure Revealed by Molecular Dynamics Simulations," J. Am. Chem. Soc. **141**, 4711–4720 (2019).
[6]J. José Segura, A. Elbourne, E. J. Wanless, G. G. Warr, K. Voïtchovsky, and R. Atkin, "Adsorbed and near surface structure of ionic liquids at a solid interface," Physical Chemistry Chemical Physics **15**, 3320–3328 (2013).
[7]C. Merlet, D. T. Limmer, M. Salanne, R. van Roij, P. A. Madden, D. Chandler, and B. Rotenberg, "The Electric Double Layer Has a Life of Its Own," J. Phys. Chem. C **118**, 18291–18298 (2014).
[8]A. A. Kornyshev and R. Qiao, "Three-Dimensional Double Layers," J. Phys. Chem. C **118**, 18285–18290 (2014).
[9]A. Elbourne, S. McDonald, K. Voïchovsky, F. Endres, G. G. Warr, and R. Atkin, "Nanostructure of the Ionic Liquid–Graphite Stern Layer," ACS Nano **9**, 7608–7620 (2015).
[10]F. Dommert, J. Schmidt, B. Qiao, Y. Zhao, C. Krekeler, L. D. Site, R. Berger, and C. Holm, "A comparative study of two classical force fields on statics and dynamics of [EMIM][BF4] investigated via molecular dynamics simulations," The Journal of Chemical Physics **129**, 224501 (2008).
[11]D. T. Limmer and A. P. Willard, "Nanoscale heterogeneity at the aqueous electrolyte–electrode interface," Chemical Physics Letters **620**, 144–150 (2015).
[12]S. W. Coles, C. Park, R. Nikam, M. Kanduč, J. Dzubiella, and B. Rotenberg, "Correlation Length in Concentrated Electrolytes: Insights from All-Atom Molecular Dynamics Simulations," The Journal of Physical Chemistry B , acs.jpcb.9b10542 (2020).
[13]A. M. Smith, A. A. Lee, and S. Perkin, "The Electrostatic Screening Length in Concentrated Electrolytes Increases with Concentration," The Journal of Physical Chemistry Letters **7**, 2157–2163 (2016).
[14]C. N. Nguyen, T. Kurtzman Young, and M. K. Gilson, "Grid inhomogeneous solvation theory: Hydration structure and thermodynamics of the miniature receptor cucurbit[7]uril," J. Chem. Phys. **137**, 973–980 (2012).
[15]J. G. Kirkwood and F. P. Buff, "The Statistical Mechanical Theory of Solutions. I," The Journal of Chemical Physics **19**, 774–777 (1951), publisher: American Institute of Physics.
[16]N. Yoshida, T. Imai, S. Phongphanphanee, A. Kovalenko, and F. Hirata, "Molecular recognition in biomolecules studied by statistical-mechanical integral-equation theory of liquids," J. Phys. Chem. B **113**, 873–886 (2009).
[17]M. C. Stumpe, N. Blinov, D. Wishart, A. Kovalenko, and V. S. Pande, "Calculation of local water densities in biological systems: A comparison of molecular dynamics simulations and the 3D-RISM-KH molecular theory of solvation," J. Phys. Chem. B **115**, 319–328 (2011).
[18]L. Ding, M. Levesque, D. Borgis, and L. Belloni, "Efficient molecular density functional theory using generalized spherical harmonics expansions," J. Chem. Phys. **147** (2017), 10.1063/1.4994281.
[19]G. Jeanmairet, B. Rotenberg, M. Levesque, D. Borgis, and M. Salanne, "A molecular density functional theory approach to electron transfer reactions," Chemical Science **10**, 2130–2143 (2019).
[20]S. Zhao, R. Ramirez, R. Vuilleumier, and D. Borgis, "Molecular density functional theory of solvation: From polar solvents to water," J. Chem. Phys. **134** (2011), 10.1063/1.3589142.
[21]G. Jeanmairet, M. Levesque, R. Vuilleumier, and D. Borgis, "Molecular density functional theory of water," J. Phys. Chem. Lett. **4**, 619–624 (2013).
[22]A. P. Lyubartsev and A. Laaksonen, "Calculation of effective interaction potentials from radial distribution functions: A reverse Monte Carlo approach," Physical Review E **52**, 3730–3737 (1995), publisher: American Physical Society.
[23]A. Chaimovich and M. S. Shell, "Coarse-graining errors and numerical optimization using a relative entropy framework," The Journal of Chemical Physics **134**, 094112 (2011), publisher: American Institute of Physics.
[24]R. Assaraf and M. Caffarel, "Zero-variance principle for monte carlo algorithms," Physical review letters **83**, 4682 (1999).
[25]J. Toulouse, R. Assaraf, and C. J. Umrigar, "Zero-variance zero-bias quantum Monte Carlo estimators of the spherically and system-averaged pair density," The Journal of Chemical Physics **126**, 244112 (2007).
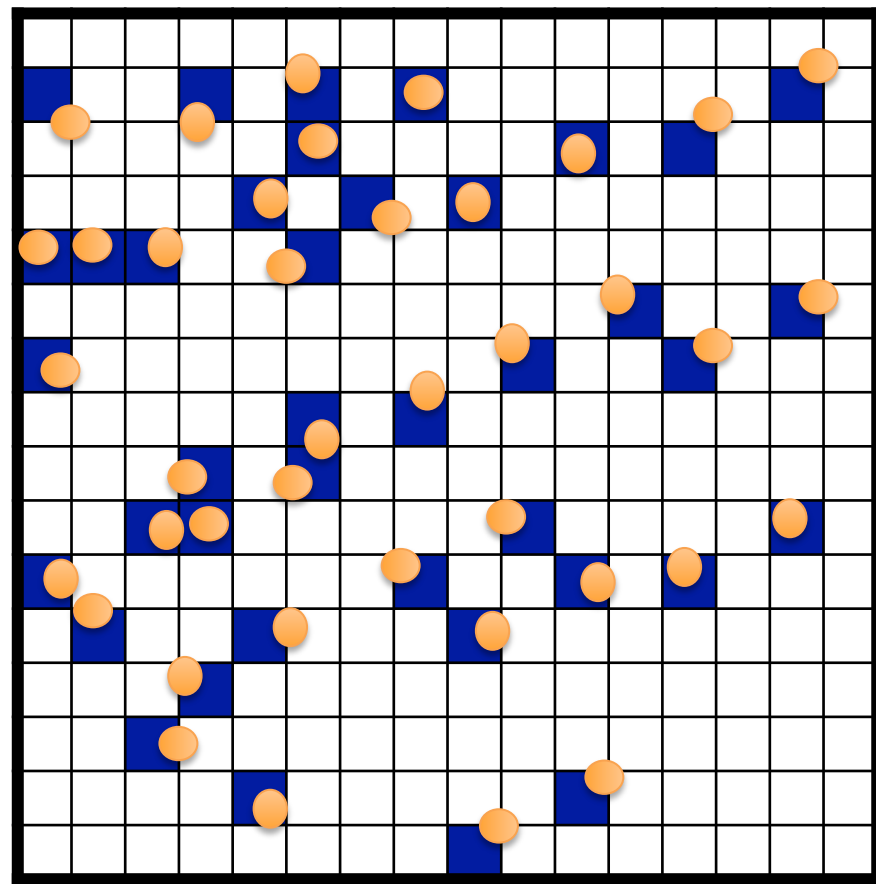
[26] R. Assaraf, M. Caffarel, and A. Scemama, "Improved Monte Carlo estimators for the one-body density," Physical Review E **75** (2007), 10.1103/PhysRevE.75.035701.

[27] A. J. Schultz, S. G. Moustafa, W. Lin, S. J. Weinstein, and D. A. Kofke, "Reformulation of Ensemble Averages via Coordinate Mapping," Journal of Chemical Theory and Computation **12**, 1491–1498 (2016).

[28] A. B. Adib and C. Jarzynski, "Unbiased estimators for spatial distribution functions of classical fluids," The Journal of Chemical Physics **122**, 014114 (2005).

[29] D. Borgis, R. Assaraf, B. Rotenberg, and R. Vuilleumier, "Computation of pair distribution functions and three-dimensional densities with a reduced variance principle," Molecular Physics **111**, 3486–3492 (2013).

[30] D. de las Heras and M. Schmidt, "Better Than Counting: Density Profiles from Force Sampling," Phys. Rev. Lett. **120**, 218001 (2018).

[31] A. Purohit, A. J. Schultz, and D. A. Kofke, "Force-sampling methods for density distributions as instances of mapped averaging," Molecular Physics **0**, 1–8 (2019).

[32] S. W. Coles, D. Borgis, R. Vuilleumier, and B. Rotenberg, "Computing three-dimensional densities from force densities improves statistical efficiency," The Journal of Chemical Physics **151**, 064124 (2019).

[33] E. Mangaud and B. Rotenberg, "Sampling mobility profiles of confined fluids with equilibrium molecular dynamics simulations," The Journal of Chemical Physics **153**, 044125 (2020), publisher: American Institute of Physics.

[34] A. J. Schultz and D. A. Kofke, "Alternatives to conventional ensemble averages for thermodynamic properties," Current Opinion in Chemical Engineering Frontiers of Chemical Engineering: Molecular Modeling, **23**, 70–76 (2019).

[35] A. Trokhymchuk, A. J. Schultz, and D. A. Kofke, "Alternative ensemble averages in molecular dynamics simulation of hard spheres," Molecular Physics , 1–20 (2019).

[36] J. L. Lebowitz and J. K. Percus, "Long-Range Correlations in a Closed System with Applications to Nonuniform Fluids," Physical Review **122**, 1675–1691 (1961).

[37] L. Belloni, "Finite-size corrections in numerical simulation of liquid water," The Journal of Chemical Physics **149**, 094111 (2018).

[38] G. Ciccotti, R. Kapral, and E. Vanden-Eijnden, "Blue Moon Sampling, Vectorial Reaction Coordinates, and Unbiased Constrained Dynamics," ChemPhysChem **6**, 1809–1814 (2005).

[39] J. E. Basner and C. Jarzynski, "Binless Estimation of the Potential of Mean Force," The Journal of Physical Chemistry B **112**, 12722–12729 (2008), publisher: American Chemical Society.

[40] C. Zhang and J. Ma, "Estimating statistical distributions using an integral identity," The Journal of Chemical Physics **136**, 204113 (2012).

[41] G. Ciccotti and M. Ferrario, "Holonomic Constraints: A Case for Statistical Mechanics of Non-Hamiltonian Systems," Computation **6**, 11 (2018).

[42] P. Krüger, S. K. Schnell, D. Bedeaux, S. Kjelstrup, T. J. H. Vlugt, and J.-M. Simon, "Kirkwood–Buff Integrals for Finite Volumes," The Journal of Physical Chemistry Letters **4**, 235–238 (2013).

[43] P. Ganguly and N. F. A. van der Vegt, "Convergence of Sampling Kirkwood–Buff Integrals of Aqueous Solutions with Molecular Dynamics Simulations," Journal of Chemical Theory and Computation **9**, 1347–1355 (2013), publisher: American Chemical Society.

[44] E. Wajnryb, A. R. Altenberger, and J. S. Dahler, "Uniqueness of the microscopic stress tensor," The Journal of chemical physics **103**, 9782–9787 (1995).

[45] T. W. Lion and R. J. Allen, "Computing the local pressure in molecular dynamics simulations," Journal of Physics: Condensed Matter **24**, 284133 (2012).

[46] Z. Zhang and W. Kob, "Revealing the three-dimensional structure of liquids using four-point correlation functions," Proceedings of the National Academy of Sciences **117**, 14032–14037 (2020).

[47] K. R. Glaesemann and L. E. Fried, "An improved thermodynamic energy estimator for path integral simulations," The Journal of Chemical Physics **116**, 5951–5955 (2002), publisher: American Institute of Physics.

[48] R. Korol, J. L. Rosa-Raíces, N. Bou-Rabee, and T. F. Miller, "Dimension-free path-integral molecular dynamics without preconditioning," The Journal of Chemical Physics **152**, 104102 (2020), publisher: American Institute of Physics.

[49] S. Fritsch, S. Poblete, C. Junghans, G. Ciccotti, L. Delle Site, and K. Kremer, "Adaptive resolution molecular dynamics simulation through coupling to an internal particle reservoir," Physical Review Letters **108**, 170602 (2012).

[50] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. v. Duijvendijk, and F. v. Wijland, "First-order dynamical phase transition in models of glasses: an approach based on ensembles of histories," Journal of Physics A: Mathematical and Theoretical **42**, 075007 (2009), publisher: IOP Publishing.
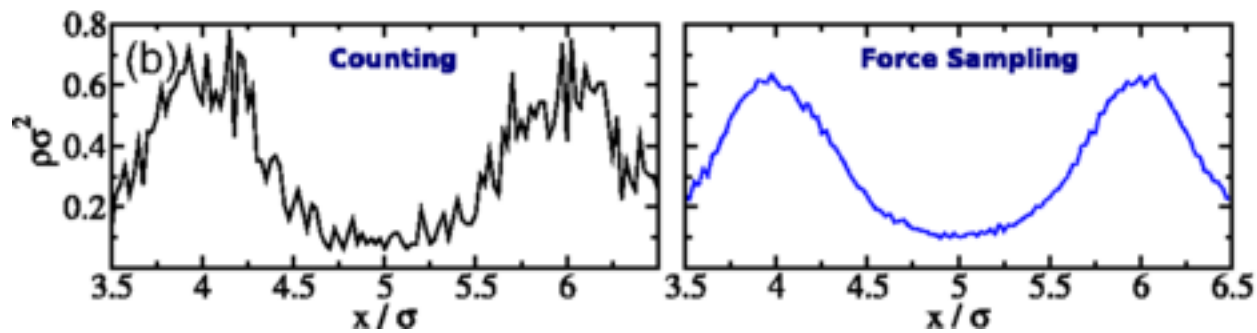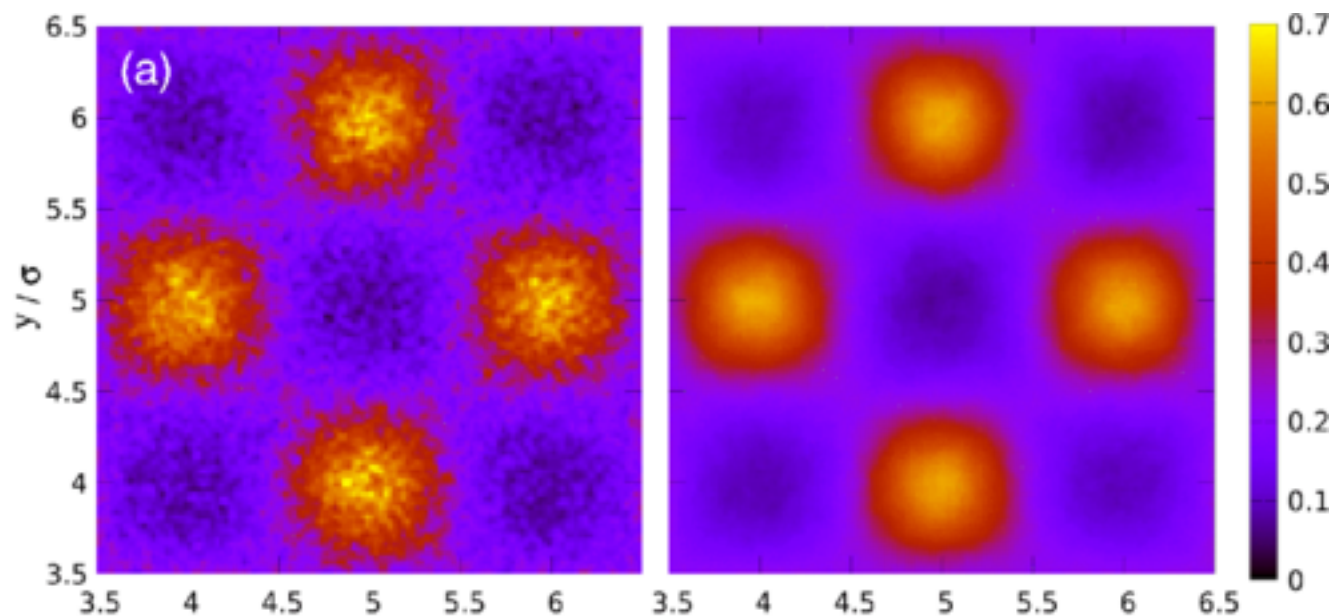
ACKNOWLEDGMENTS

(a) h

(b) $h' = h/2$

(a) (b) z

(c) $h = 0.2 \ \mathring{A}$ (d) $h' = h/3$

$P_z(\mathbf{r})/\,\mathrm{D}\cdot\mathring{A}^{-3}$

Histogram
Force

$z/\mathring{A}$

	Histogram Method	Force Method
ρ		
ρ_{q}		

(a)

Histogram ($\Delta r = 0.005\sigma$)
Converged after 10^4 steps
Force sampling

(b)

Histogram ($\Delta r = 0.005\sigma$)
Histogram ($\Delta r = 0.01\sigma$)
Force sampling

$v(r)$

r (σ)

(a)

(b)

(c)