Comparison of Umbrella Sampling Methods and Steered Molecular Dynamics for Computing Free Energy Profiles of Toluene Molecules through Phospholipid Bilayers

Sang Young Noh* and R. Notman†

Department of Chemistry, University of Warwick

(Dated: August 16, 2019)

The Umbrella Sampling method (US) has been the standard for use in the computation of free energy (FE) profiles of reaction coordinates. However, in practise, the results from US methods are dependent on the strength of the biasing potential and the spacing of windows along the reaction coordinate, which can make this approach challenging to implement in practice, due to the challenges of finding the right level of overlap between the sampling bins and finding the correct spring potentials that may be required across the reaction coordinate bins. An alternative approach, using the Steered Molecular Dynamics implementation of the Jarzynski Equality (JE-SMD) has been identified as a viable alternative to equilibrium sampling methods for measuring the FE change across a reaction coordinate, where the main advantage over US comes from the evaluation of equilibrium FE values from the average of rapid non-equilibrium trajectories, leading to a much more rapid convergence towards the FE profile and avoiding the bin-positioning issues that comes with the US method. Here, we take multiple corrections of the JE-SMD method to measure the FE change for the translocation of a toluene molecule across a lipid bilayer, and measure its accuracy and computational efficiency with the US simulation as the benchmark for the FE profile to converge towards. This system serves as a rapid benchmark for how efficient the non-equilibrium method would fare in larger and more complex systems, as well as allow the streamlining of the computation of FE profiles of the transport of drug-like molecules across a model lipid bilayer to estimate the partition coefficient of drug-like molecules. From the comparison of the methods, we show that there are sampling issues regarding the JE-SMD method, which may be alleviated with a slower velocity, but at the cost of computational efficiency. We observe the most significant correlation to the US result seen with the raw JE-SMD formulation. From this study, we assess that the US method is still the most viable approach for computing the FE profile.

* S.Y.Noh@Warwick.ac.uk

† R.Notman@Warwick.ac.uk

I. INTRODUCTION

The study of the translocation of molecules through lipid bilayers has been a subject of much interest, due to it being the first barrier of entry inside the cell. Determining the ease of entry for a molecule may therefore correlate directly with its effectiveness as a drug-delivery molecule. The factors that affect the mechanism of entry through the bilayer depends on a number of factors, including the chemical composition of the head and tailgroups, headgroup area, tailgroup composition, and the hydrophilicity/hydrophobicity of the chemical moieties composing the bilayer and the drug molecule. Depending on the property of the drug molecule and the local membrane properties, we may observe passive diffusion like processes, or energy-driven (ATP) endocytosis-type mechanisms through membrane proteins [1, 2]. The use of *molecular simulation* to derive these values has been increasing, due to the increasingly accurate parameterization of forcefields (FF) representing the drug-like molecules and the variety of lipids. FFs parameterizations such as CHARMM, AMBER, MARTINI and ELBA show a variety of resolutions for the molecules [3–6]. The main advantage of simulations methods for studying these types of systems is that they can be set and adjusted with fixed parameters (such as pressure, volume, chemical potential). From this, it is possible to use the tools of statistical mechanics to compute the state variables, which may be difficult to compute experimentally. In particular, the *free energy* (FE) of processes has become an essential property to compute, as it dictates spontaneous nature of the permeation, the energetic cost, and can be used to compute the *partition coefficient* of the drug molecule, which may be measured experimentally [7, 8].

A number of enhanced sampling methods have been used to determine the FE profile across a reaction coordinate in simulations. In particular, the *umbrella sampling* (US) method has been ubiquitous in its use, due to its intuitive nature and relative simplicity of implementation [9–11]. In the US method, the reaction coordinate is divided into bins, and within each bin the reaction coordinates are held by a Hook's potential, allowing the sampling of the fluctuating reaction coordinates. The resulting sampling histograms are then recombined to reconstruct the FE profile. The weaknesses of the US method comes from its high computational cost - the high energy parts of the phase space needs to be sampled adequately with the right overlap between each sampling space, and the discrete sampling each bin of the reaction coordinate requires a large trial and error process to find parts that may require additional sampling bins. Alternatively, Jarzynski [12] demonstrated a revolutionary equality (known in literature as *Jarzynski's equality* (JE)), which showed that independent of the velocity of the process (hence, can be a non-equilibrium process), the *force-distance* curve along the reaction coordinate can be used to compute the work, a path-dependent function, to compute the FE of the process, an *equilibrium* state function.

The computation of the FE using JE based methods introduced *fast-switching* events (a series of high constant velocity simulations, relative to analogous US simulations) as a possible efficient alternative to the US method. Should the validity of the equality be widely applicable, the advantage of the method compared to the US method is obvious - one would not need to sample discrete bins along the reaction coordinate, and would therefore computational resources which would enable computation of more complex reaction coordinates, as we would not explicitly require the system to equilibrate before extracting the FE value. By adapting a moving Hook's potential, Schulten *et al* suggested the *steered molecular dynamics* method [13] which can constrain the particle/molecule along the reaction coordinate and compute the work distribution that

is required for the calculation of the FE (Hence, we abbreviate the practical implementation of JE as JE-SMD).

Previous studies comparing the US and JE-SMD have established that the JE-SMD method can give comparable results to the US method for an accurate estimate of the PMF, depending on the velocity chosen for the moving harmonic potential [13]. However, evidence has also arisen which also suggests that sampling issues arise when implementing the method to complex many body systems. Verification of the JE involve relatively simple processes, such as the extension of a DNA residue [13] and the unfolding of RNA hairpins [14], which does not therefore, dismiss the possibility of error in the JE-SMD method when applied to many body systems. The trial systems which was used to verify the JE method used a Langevin dynamics parameter to mimic thermal fluctuations, and it's small size allowed the repeated simulations (of the order 10,000 runs and more) to be carried out with ease. This would not be feasible for the vast majority of complex systems, where the simulation costs relating to the level of atomic detail and size are the limiting factors for making comparable number of simulations. The application of the JE-SMD in larger simulations (of 10 - 100 nm scale) have shown mixed results; Kuyucak *et al* [15, 16] showed that the JE-SMD FE measurements in lipid-bilayer based systems, such as the flow of water molecules across a carbon nanotube and a gramicidin A/bilayer channel showed a large discrepancy compared to the benchmark US results, which was suggested to be affected due to the slow relaxation of the surrounding system, which is dependent on the pulling velocity. This shows that the JE-SMD method considerable promise as a FE sampling method that is comparable to the US method, but requires the right optimization in the interpretation of the JE-SMD work distributions.

In this article, we present a systematic study of the enhanced sampling FE trajectories of a toluene molecule translocating through a model 1,2-Dioleoyl-sn-glycero-3-phosphocholine (DOPC) lipid bilayer using the US and the JE-SMD method. In particular, we concentrate on the effects of modifying the velocity and the JE-SMD interpretation that has been suggested in literature. We look at the convergence behaviour of the JE-SMD method towards the US result with the assumption that the FE profile from the US method represents the true FE profile, and we analyse its computational efficiency.

II. SIMULATED SYSTEMS

A. Forcefield Parameters

The *ELBA* biomolecule *mixed resolution* coarse-grained and all-atomic (CG-AA) compatible FFs was provided from the research group of Mario Orsi and Johnathan Essex [5, 17]. The ELBA-CG-AA hybrid model offers an intriguing possibility for new insights, with AA-level detail coupled with CG-level simulation scales. The AA-CG compatibility of ELBA allows an arsenal of explicit FF potentials to be used in combination with the CG-lipid FF; the AA interatomic potentials of the GAFF (General Amber FF) can be combined with the CG potentials of the lipid molecules [18]. Hence, the model allows inter-compatibility with the traditional AMBER FF (consistent AM1-BCC charge models with identical VdW parameters) which gives it wide applicability to known libraries of small molecules. The details of the ELBA FF have been included in the supplementary information.

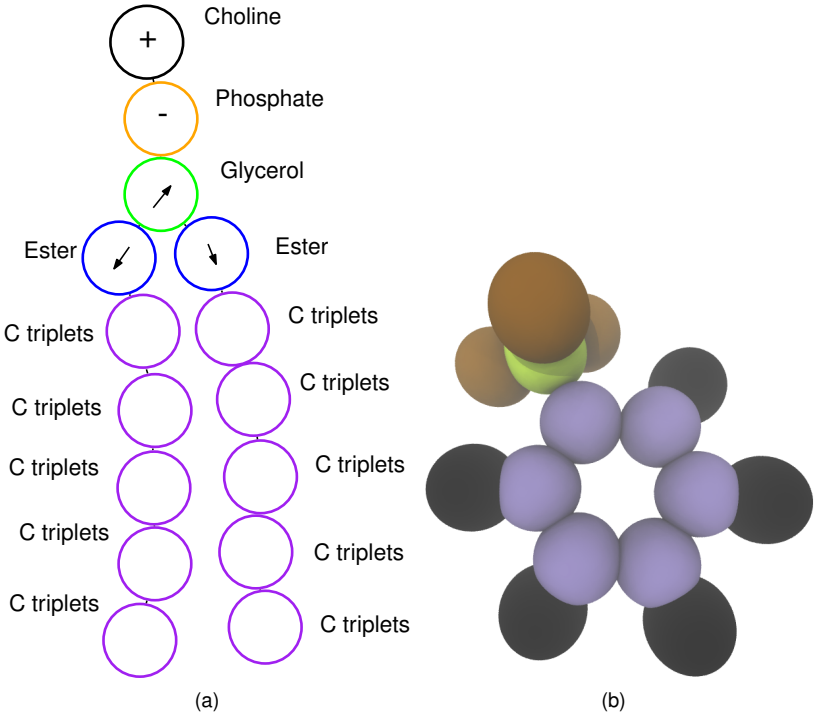


FIG. 1: (a) a schematic of the CG DOPC molecule and the toluene molecule, where the headgroup is represented by the choline and phosphate beads (both charged), the glycerol and ester region is represented by beads with opposing point dipoles, while the hydrophobic tailgroups are represented by C-triplets, where each bead represents a triplets of carbon atoms (b) shows the toluene molecule, which is built from the AMBER GAFF forcefield [18]. The graphical illustration shown in (b) was produced using VMD [19].

III. SIMULATION PARAMETERS

Each simulation was run with the *LAMMPS* molecular dynamics package [20]. The range of systems was used to determine the effect of composition to the PMF of the toluene translocation. Each bilayer system was initially run on an *NVT* run with the Langevin thermostat for 100 ns at temperatures of 298 K. A production simulation run with the Berendson barostat and Langevin thermostat (i.e. a *NPT* equivalent) was then carried out for 100 ns to ensure that the system has been equilibrated. The temperature and pressure of the *NPT* simulation was set to 298 K and 1 atm respectively. The pair-wise interactions used a shifted-force type with a 1.2 nm cut-off. The rRESPA (reversible-Reference System Propagator Algorithm) was used to divide the total timestep to compute the AA and CG components on a different timescales [21]. Outer level timesteps updated the CG-CG interactions (i.e. between the ELBA-beads) while the AA-CG and AA-AA interaction were updated in the inner level. The inner timestep was computed every 1.0 fs, while the outer timestep was computed every 4.0 fs. The system contains 128 DOPC molecules in total, with 64 lipids on the top and bottom leaflets. 4232 water beads in total were placed above and below the leaflets. The input script has been listed in the supplementary information (Listing 1). Figure 1 shows the schematic of the CG-DOPC molecule and the molecular structure of toluene used for this simulation.

A. Free Energy Calculations

1. Umbrella Sampling

For the US simulations, the reaction coordinate (the z coordinate across the bilayer normal) was divided in 1 Å intervals from the center of the bilayer to the headgroup region of the leaflet. The bias potentials of each US window were set to 2.5 kcal mol^{-1}, increased to 5.0 kcal mol^{-1} for 2 Å above and below the headgroup region in the z coordinate (this corresponds to a FE barrier region and follows closely the US procedure followed by Genheden *et al* [7]). The benchmarking US simulation for the toluene molecule was also undertaken under similar conditions - the toluene molecule was tethered at 1 Å intervals apart ranging from 0 Å (bilayer center) to 30 Å (above the headgroup region). In total, 30 US windows were equilibrated with an initial run of 10 ns, and production window for *weighted histogram analysis method* (WHAM) [22] analysis were run for 30 ns for each bin. Further detail on the WHAM procedure is included in the supplementary information.

2. Jarzynski Equality

The Jarzynski equality (JE) is defined as:

$$e^{-\beta \Delta G(z)} = \langle e^{-\beta W(z)} \rangle \tag{1}$$

where β is $\frac{1}{k_\beta T}$, where k_β is the Boltzmann constant, T is the temperature, ΔG(z) the FE at the given *z* coordinate, and W the work [12]. What is intriguing with this equality is that W is a path-dependent property, whereas ΔG is a state function - an equilibrium property. Hence, this value equates a non-equilibrium property to a equilibrium property. Inherent issues regarding the practical use of the JE is the inherent *bias* related to insufficient sampling and the dissipated work. A direct interpretation of the JE to map a FE profile can be interpreted from the following equation:

$$\Delta G_{\mathsf{J}} = \frac{1}{\beta} \ln \left[\frac{1}{N} \sum_{i}^{N} e^{\beta W_i(z)} \right] \tag{2}$$

Where N is the total number of work samples we are working with. The use of the JE for FE calculations may suffer from a significant amount of bias (i.e. the difference between the expected value of the FE and its estimate, ΔG - ΔG_{J}) due to insufficient sampling along the reaction coordinate. Also, the exponential average value ($\frac{1}{N} \sum_i^N e^{\beta W_i(z)}$) is dominated by rarely occurring small work values. The second order cumulant expansion term for the JE was used to correct this bias due to this sampling problem [13]:

$$\Delta G_{\text{cumulant}} = \langle W(z) \rangle - \frac{N}{N-1} \frac{\beta}{2} (\langle W(z)^2 \rangle - \langle W(z) \rangle^2) \tag{3}$$

Where the ($\langle W(z)^2\rangle - \langle W(z)\rangle^2$) term is the variance of the work along the reaction coordinate (where the angular brackets represent the averaged value over *N* trajectories), and the $\langle W(z)\rangle$ term is the averaged work - this modified JE term for the FE difference is valid on the condition that the work distribution along the reaction coordinate is *Gaussian*, as this enables the elimination of cumulants higher than that of second order of the JE to equal 0. When using this estimator to compute the FE, another factor to consider is the distribution of the dissipation work, where the dissipated work is defined as the difference between the averaged work and the true FE difference at the reaction coordinate - the magnitude of the dissipated work dictates the width of the Gaussian probability distribution. The dissipation work may therefore be estimated by:

$$W(z)_{\text{dissipation}} = \frac{1}{2}\beta\sigma^2_{W(z)} \tag{4}$$

Where $\sigma^2_{W(z)}$ represents the variance of the work. We note that this term is identical to the factor in variance of the work in the term $\frac{\beta}{2}(\langle W(z)^2\rangle - \langle W(z)\rangle^2)$, which shows that the variance of the work along the RC is equal to the $W(z)_{\text{dissipation}}$. The probability of observing a trajectory with negative dissipation work [23] can be described as:

$$P(W(z)_{\text{dissipation}} < 0) = \frac{1}{2}[1 - \text{erf}(\sqrt{\langle W(z)_{\text{dissipation}}\rangle/2})] \tag{5}$$

Where $\langle W(z)_{\text{dissipation}}\rangle$ represents the averaged dissipation work along the trajectory. The equation implies that with a larger $W_{\text{dissipation}}$, the probability of observing a negative dissipation work sharply decreases. Hence, an increased magnitude of dissipated work corresponds to a lower probability of observing negative work events, and is undesired. An alternate method for taking into account the bias of insufficient sampling was suggested by Gore *et al* [23, 24] who proposed an alternative correction to reduce the sampling bias, following the observation that the bias over a small number of trajectories ($N < 10^2$) has a linear relationship with N, meaning that an approximate bias can be numerically computed. This modified JE sampling bias is defined as:

$$B_J = \frac{\langle W(z)_{\text{dissipation}}\rangle}{N^{\alpha_b}} \tag{6}$$

α_b is the rate at which the bias goes to zero with a small number of sampling:

$$\alpha_b = \frac{\ln[\beta C_b \langle W(z)_{\text{dissipation}}\rangle]}{\ln[C_b(e^{2\beta\langle W(z)_{\text{dissipation}}\rangle} - 1)]} \tag{7}$$

where C_b is a fitted parameter that determines the boundary of small N - large N regime. We have utilised C_b = 15, as tested by Gore [23], which defines the intersection point for the small N/large N regime. From this, we can estimate a bias-corrected Jarzynski estimator:

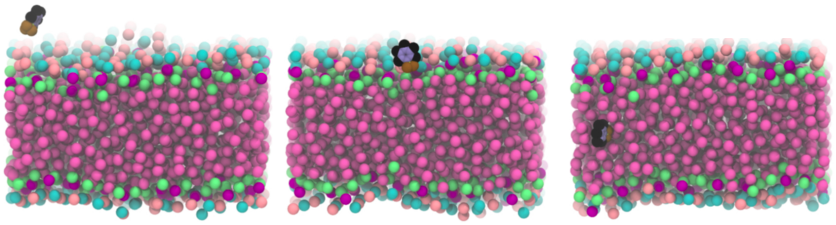
$$\Delta G_{\text{bias}} = \Delta G_{\text{J}} - B_J \tag{8}$$

When computing average work along the reaction coordinate, we followed a block-averaging procedure, where every value in an 1 Å interval along the reaction coordinate was binned and averaged, and the variance within each bin was determined using the mean value of each bin. For each bin, we computed several terms to assess the validity of the JE for the elucidation of an accurate FE profile. We computed the ΔG_{J} term, $\Delta G_{\text{cumulant}}$, ΔG_{bias}, and the $\sigma^2_{\text{W(z)}}$. The JE-SMD to compute the work procedure was implemented as follows: after 10 ns of equilibration, the toluene molecule was inserted held at 30 Å from the bilayer center of mass, in the normal (z) direction.

For the JE-SMD simulations, four different constant velocities were assigned - at *constant velocities* of 8 $\times$ 10^{-6} Å fs^{-1} (*faster*), 5 $\times$ 10^{-6} Å fs^{-1} (*fast*), and 8 $\times$ 10^{-7} Å fs^{-1} (*slow*) and 5 $\times$ 10^{-7} Å fs^{-1} (*slower*). For each simulation, the toluene was pulled from z coordinate values from 30 Å to 0 Å. These velocities are consistent with those used in previous studies and slower velocities were also included here to systematically analyse the effect of velocity. For example, Schulten *et al* used velocities of 0.1 Å ns^{-1} (0.0001 Å fs^{-1}) , 10 Å ns^{-1} (0.01 Å fs^{-1}), 100 Å ns^{-1} (0.1 Å fs^{-1}) when validating JE-SMD through stretching deca-alanine [13], where in each case highly exaggerated velocities were used to test the validity of JE in a truly non-equilibrium environment. In other examples, Kucuyak *et al* [15, 16] used velocities ranging from 1.25 to 10 Å ns^{-1} (Hence, 1.25 $\times$ 10^{-6} Å fs^{-1} to 10.0 $\times$ 10^{-6} Å fs^{-1}), which are directly comparable to the velocities used with our simulations. Given this, we can conclude that the velocities we have chosen are reasonable to test the JE-SMD method. The *faster* and *fast* velocities can be defined as the 'fast-growth' regime, while *slow* and *slower* velocities can be recognised as the 'slow-growth' regime. The biasing potential (Hooks potential) to anchor the toluene molecule to the reaction coordinate used a force constant of 25 kcal mol^{-1}. The work values were collated over 20 repetitions of the pulling simulation. To analyse the effect of using increased sampling blocks, we measured the ΔG_{J} term, $\Delta G_{\text{cumulant}}$, ΔG_{bias} and the $\sigma^2_{\text{W(z)}}$ term over of N = 10, 15 and 20 simulation runs. Table I lists the scope of the simulations undertaken in this study.

Index	**Method**	N	v **(Å fs^{-1})**	k **(kcal mol^{-1} Å)**	t **(ns)**
1	JE-SMD	20	8 x 10^{-6}	25	100
2			5 x 10^{-6}		120
3			8 x 10^{-7}		720
4			5 x 10^{-7}		1200
5	US	30	-	25/50	900

TABLE I: List of simulations run in this study - including the JE-SMD simulations at different velocities, and the US simulations used as the benchmark/control to compare against. N represents the number of simulation trajectories used for averaging the work/sampling bins, v represents the velocity of the toluene in the JE-SMD simulations, k represents the spring constant used for restraining the toluene molecule to the reaction coordinate, and t represents the total simulation time.



(a) z = 30 Å (b) z = 20 Å (c) z = 0 Å

FIG. 2: Trajectory snapshots for the JE-SMD/US simulations for the toluene/DOPC, showing the z coordinate of the toluene molecule in each snapshot. The graphical illustrations were produced using VMD [19].

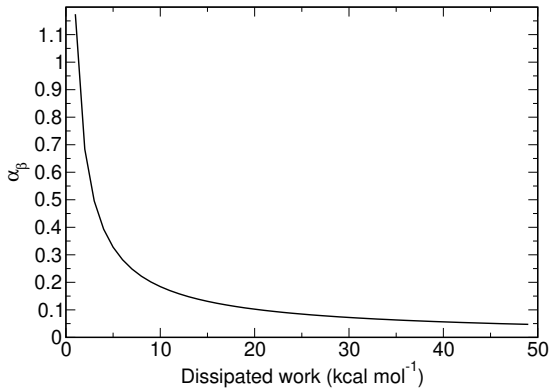


FIG. 3: The plot for $\alpha_b = \frac{\ln[\beta C_b \langle W(z)_{\text{dissipation}} \rangle]}{\ln[C_b(e^{2\beta \langle W(z)_{\text{dissipation}} \rangle} - 1)]}$, for dissipation work between 0 - 50 kcal mol^{-1}. Interpreting the $\frac{\beta}{2}\sigma_{\text{W}}^2$ as the dissipation work, the α_b was selected when computing the $\Delta G(z)_{\text{bias}}$ from the JE-SMD simulations.

IV. RESULTS

A. JE-SMD/US Comparison

Figure 2 shows the snapshots of the JE-SMD and US trajectories of the toluene molecule being pulled through the DOPC bilayer. Table II shows the compiled ΔG values and growth regimes for each velocity of the JE-SMD implementations at

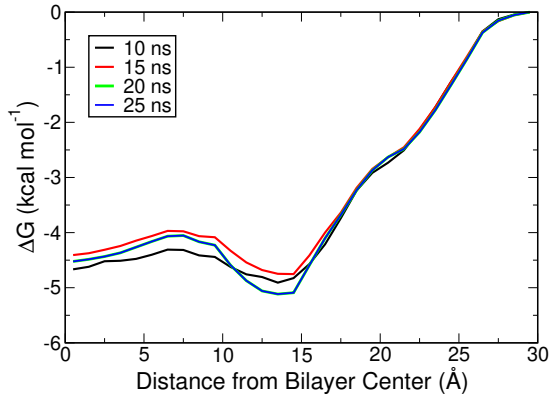


FIG. 4: The US simulations for the toluene. The WHAM collated values at 5, 10, 15 and 25 ns is shown to illustrate the convergence of the FE profile at approximately 30 ns.

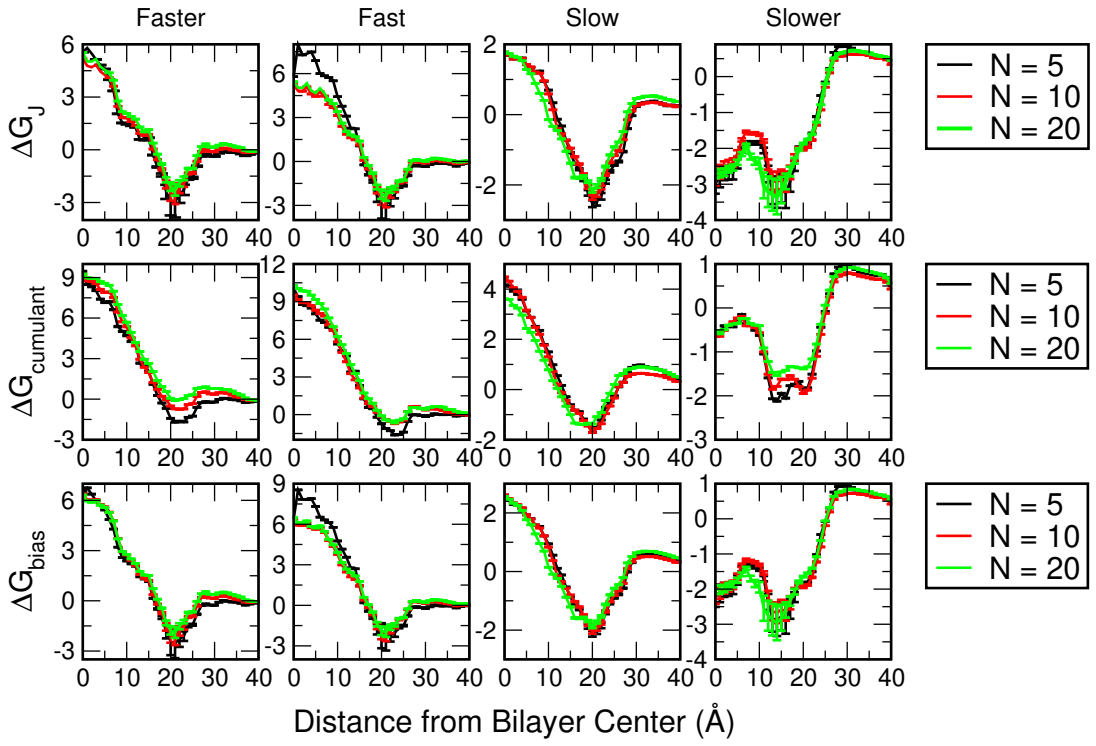


FIG. 5: The FE profiles computed from the v = 8 $\times$ 10^{-6} Å fs^{-1} (faster), v = 5 $\times$ 10^{-6} Å fs^{-1} (fast), v = 8 $\times$ 10^{-7} Å fs^{-1} (slow), and v = 5 $\times$ 10^{-7} Å fs^{-1} (slower) JE-SMD simulations. The first row shows the ΔG_{J}, the second row shows the $\Delta G_{\text{cumulant}}$, and the third row shows the ΔG_{bias}. For each row, the convergence of JE-SMD simulation profiles was shown from N = 10, 15 and 20 simulations respectively. The units for ΔG in each case was kcal mol^{-1}.

four different velocities. Figure 3 shows the α_b plot used for computing the ΔG_{bias} values. Figure 4 shows the benchmark US simulation with a FE change of -5.2 kcal mol^{-1} (the sampling profile is shown in Figure S1) (corresponding to 23.012 kJ mol^{-1}, which is comparable to the values obtained in this system (BERGER AA (-12.5 kJ mol^{-1}) and MARTINI CG FF values (-20.92 kJ mol^{-1})) [25–27]. Here, the key feature to compare with the JE-SMD results is the energy trough near trough near z = 15 Å. Figure 5 shows the ΔG_{J}, $\Delta G_{\text{cumulant}}$, and ΔG_{bias} for each velocity of the JE-SMD simulations. Each JE-SMD plot also shows convergence as averaging is carried out over successively larger numbers of JE-SMD simulations. For the v = 8 $\times$ 10^{-7} Å fs^{-1} (*faster*) simulations, we see an overall change of -3.0 - 6.0 kcal mol^{-1}, -3.0 - 9.0 kcal mol^{-1} and -3.0 - 6.5 kcal mol^{-1} for the ΔG_{J}, $\Delta G_{\text{cumulant}}$ and ΔG_{bias} values respectively. Similar results are seen with the v = 5 $\times$ 10^{-7} Å fs^{-1} (*fast*) simulations, which show very similar patterns with increasing N simulations - it is clear that the JE-SMD simulations in the 'fast-growth' regime (faster and fast velocities), does not reproduce the US results effectively, and that there is minimal effect on the FE profile upon increasing the number of simulations, N.

Interpreting the 'slow-growth' regime simulations, for the v = 8 $\times$ 10^{-7} Å fs^{-1} JE-SMD simulations (*slow*) , we see a overall improvement in the PMF profiles - we see an overall change of -2.0 - 2.0 kcal mol^{-1}, -2.0 - 4.0 kcal mol^{-1} and -2.0 - 2.5 kcal mol^{-1} for the ΔG_{J}, $\Delta G_{\text{cumulant}}$ and ΔG_{bias}. There is an 0 and 20.0 Å with the $\Delta G_{\text{cumulant}}$ and ΔG_{bias}, which represents a marginal improvement, but clearly not close to convergence to the US result. At velocities of v = 5 $\times$ 10^{-7} Å fs^{-1} (*slower*), we see the closest consistency with the US profile - at N = 10 simulations, with the emergence of a FE

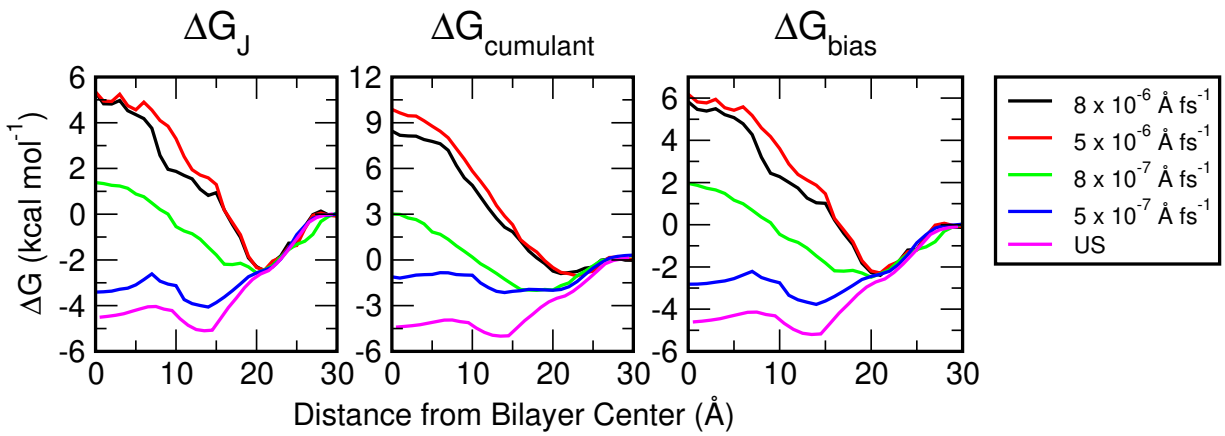


FIG. 6: The collated FE profile JE-SMD simulations, as shown by the v = 8 $\times$ 10^{-6} Å fs^{-1}, 5 $\times$ 10^{-6} Å fs^{-1}, 8 $\times$ 10^{-7} Å fs^{-1}, 5 $\times$ 10^{-7} Å fs^{-1} (corresponding to the faster, fast, slow, and slower velocities respectively) JE-SMD simulations (ΔG_{J}, $\Delta G_{\text{cumulant}}$, ΔG_{bias}) with the US result as the benchmark profile.

minimum at ~15 Å. Within these profiles, we see the key features that are consistent with the US profile - the trough at 10.0 - 20.0 Å and 0.0 - 7.0 Å. This problem is marginally improved by averaging over larger number of samples (N = 10, 20), where we see a overall FE change of -3.2 kcal mol^{-1}, -0.5 kcal mol^{-1} and -2.5 kcal mol^{-1} for the ΔG_{J}, $\Delta G_{\text{cumulant}}$ and ΔG_{bias} profiles respectively. Here, it is clear that the ΔG_{J} and ΔG_{bias} shows the clearest consistency with the US results, while the $\Delta G_{\text{cumulant}}$ profile shows a smaller FE minima near the 10 - 20 Å region. The $W(z)_{\text{dissipation}}$ is shown in the supplementary information (Figure S2). We observe that a dramatic trough below from z = 20 Å, which represent the bilayer region. The trend we see is an overall decrease in the magnitude of the $W(z)_{\text{dissipation}}$, where we see convergence towards values of 10.0 kcal mol^{-1} (*faster*), 8.0 kcal mol^{-1} (*fast*), 5.5 kcal mol^{-1} (*slow*), 4.5 kcal mol^{-1} (*slower*).

FE Method	v **(Å** fs^{-1}**)**	**ΔG (kcal** mol^{-1}**)**	**type**
ΔG_{J}	8 $\times$ 10^{-6}	6.0	fast-growth
	5 $\times$ 10^{-6}	5.5	fast-growth
	8 $\times$ 10^{-7}	2.0	slow-growth
	5 $\times$ 10^{-7}	-3.2	slow-growth
$\Delta G_{\text{cumulant}}$	8 $\times$ 10^{-6}	9.0	fast-growth
	5 $\times$ 10^{-6}	9.0	fast-growth
	8 $\times$ 10^{-7}	3.8	slow-growth
	5 $\times$ 10^{-7}	-0.5	slow-growth
ΔG_{bias}	8 $\times$ 10^{-6}	6.0	fast-growth
	5 $\times$ 10^{-6}	6.0	fast-growth
	8 $\times$ 10^{-7}	2.5	slow-growth
	5 $\times$ 10^{-7}	-2.5	slow-growth
US	-	-5.2	-

TABLE II: Data for the overall change in ΔG for each JE-SMD and the benchmark US simulation.

V. DISCUSSION

From the work distributions along the reaction coordinate (Figure S3), we have shown that the distribution follows a series of Gaussian-like profiles, which satisfies the condition for the cumulant approximation of the JE-SMD method to be applicable. In the JE-SMD validation by Schulten [13], they reported the $W(z)_{dissipation}$ to be between 1.9 to 4.3 kcal mol^{-1} (3.1 k_BT - 7.1 k_BT), where a smaller estimate corresponded to a PMF profile that was closer in convergence with the corresponding US example. With our toluene simulations, we observe $W(z)_{dissipation}$ values in the ranges of 1.0 - 10.0 kcal mol^{-1}, 1.0 - 8.0 kcal mol^{-1}, 1.0 - 5.0 kcal mol^{-1}, and 1.0 - 4.5 kcal mol^{-1} for the *faster*, *fast*, *slow*, *slower* simulations respectively (Figure S2). The highest values for $W(z)_{dissipation}$ occurs in the bilayer interior ($z < 20$ Å). A greater value of $W(z)_{dissipation}$ results in a broader probability distribution for the $W(z)_{dissipation}$ [23, 24], and shows that the convergence of an accurate estimate of $P(W(z)_{dissipation})$ is a concern within the bilayer interior regions. Other bilayer comparison studies of the JE-SMD and US methods by Kuyucak *et al* [15, 16] show that even a small inclusion of Coulombic interactions and flexible peptide molecules can affect efficient sampling with both fast and slow simulations due to the slow relaxation time. In other studies, Warshel *et al* [28] has shown that, showing that the convergence of a simple gramacidin A channel takes of the order of s, which clearly highlights this issue. Hence to make a valid comparison of the FE methods, we were compelled to select a system that was minimally affected by these factors. In our example, the toluene molecule was explicitly chosen as a rigid, non-ionic, hydrophobic molecule which would circumvent the sampling issues related to the Coulombic potentials.

The final comparison between the JE-SMD and US FE profiles are shown in Figure 6. In the *faster*, *fast* and *slow* cases, the regions of 0 - 20 Å fails to capture the overall shape or magnitude entirely. It is only with the 5×10^{-7} Å fs^{-1} (*slower*) JE-SMD samples that we see a pattern of consistency with the US result in terms of the shape of the profile, though not with the magnitude. The closest estimate is the change of -2.5 kcal mol^{-1} with the ΔG_{J}, with minor changes seen with the ΔG_{bias} result. The $\Delta G_{\text{cumulant}}$ result shows the least convergence towards the US result, where the small trough in ranges of 10 - 15 Å is barely noticeable. The cumulant estimator, which was designed to correct the original JE-SMD estimator, proves inferior in this case compared to both the original and bias-corrected estimator. The overall trend we observe with a JE-SMD simulation is that with a slower velocity, the FE profile for the JE-SMD shows marginal improvements in its profile, and the magnitude of change becomes smaller (i.e approaches the US). Therefore, Schulten *et al* [13] has commented that the JE is valid for processes at any velocity, even at velocities where it becomes seemingly more practical compared to the US method, the bias due to the under sampling of rare-trajectories makes the JE-SMD method impractical to use unless when used for the simplest systems. The fact that the energy minimum is not placed within the center of the DOPC bilayer in the FE profile for both the US and JE-SMD seems to indicate a general sampling issue in both methods. This issue may be an issue on computing the correct membrane binding and partition coeffients.

VI. CONCLUSION

We have used a model toluene/bilayer system to establish the validity of JE-SMD method compared to the US method for computing the FE profile of a molecule crossing of membrane. We have divided the JE-SMD experiments into two regimes of velocities - a 'fast-growth' regime where the total simulation time is low compared to the US method, and a 'slow-growth' regime where the total simulation time is comparable or exceeds the total simulation for US sampling. The PMF profile from the US simulations confirmed that the FF model we have used with this simulation can be trusted to produce consistent PMF data with known literature, and hence is valid to use as a benchmark to compare the JE-SMD simulations. To measure the efficiency and accuracy of the JE-SMD method to one that is comparable to the US method, we implemented three interpretations of the JE - the raw JE interpreter, ΔG_J, the cumulant second-order interpreter $\Delta G_{cumulant}$ which corrects the sampling bias, and ΔG_{bias}, which is an alternate method for taking into account high sampling bias. Within the scope of a simple bilayer/toluene simulation, we failed to see a convergence towards US result for $N = 20$ repeat runs in the 'fast-growth' regime, with all interpretations of the JE-SMD, while in the 'slow-growth' regime, results varied; while we observed significant improvements in terms of the appearance of peaks and troughs in the FE profile, the JE-SMD did not fully converge towards the US result at the slowest velocities, and each bias corrected interpretations of the JE-SMD did not significantly alleviate the sampling problem.

It is clear that the primary factor in improving the JE-SMD profile is the convergence of the $W(z)_{dissipation}$ and sampling - slower velocities allows for this, but at the rate which this sampling space is reached, the computational cost and number of simulations required becomes prohibitive, and hence, inefficient when compared to the US method - In the 'slow-growth' regime, we have already allocated a longer total simulation time compared to the US simulations, which again indicates that the the JE-SMD method needs modification if it is to be practically utilised. As this bilayer system was designed to be a benchmark for larger systems with inhomogenous components, it raises questions into how this sampling issue seen in this study would apply to systems with heterogenous and complex components. Hence, while it is clear that the JE-SMD method provides a path towards an efficient alternative to FE sampling, issues remain in its interpretation, and further work is required to correctly account for significant bias of the region. Recently, alternative modifications using *adaptive stochastic perturbation protocols* (ASPP) [29] or *multistep trajectory combination* (MSTC) [30] have been suggested for improvements onto the JE. In the case of ASPP method, the method improves the computed FE profile by the widening of the work distribution, whilst in the case of the MSTC method, the small numbers of trajectories combined in steps. Both these methods require further investigation, and will be explored in a future work.

VII. ACKNOWLEDGEMENTS

This work was supported by the EPSRC and the Center for Scientific Computing (CSC) of the University of Warwick. Computing resources was provided by the University of Warwick Tinis/Minerva clusters. I would like to thank Rebecca

Notman for the assistance and guidance in the completion of this work.

[1] M. Edidin, “Lipids on the frontier: a century of cell-membrane bilayers,” Nature Reviews Molecular Cell Biology **4**, 414 – 418 (2003).

[2] D. E. Engelman, “Membranes are more mosaic than fluid,” Nature **438**, 578 – 580 (2005).

[3] B. R. Brooks, C. L. Brooks, A. D. Mackerall, L. Nilson, R. J. Peterlla, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Carflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazardis, J. Ma, V. O. aand E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venble, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus, “CHARMM: The Biomolecular simulation program,” Journal of Computational Chemistry **30**, 1545 – 1615 (2009).

[4] S. J. Marrink, H. J. Risselada, S. Yefimov, P. D. Tieleman, and A. H. de Vries, “The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations,” Journal of Physical Chemistry B **111**, 7812–7824 (2007).

[5] M. Orsi and J. W. Essex, “The ELBA Force Field for Coarse-Grain Modeling of Lipid Membranes,” Plos One **6**, e28637 (2011).

[6] M. Orsi, “Comparative assessment of the ELBA coarse-grained model for water,” Molecular Physics **112**, 1566 – 1576 (2014).

[7] S. Genheden and J. W. Essex, Journal of Chemical Theory and Computation **11**, 4749 – 4759 (2015).

[8] K. Bittermann, S. Spycher, S. Endo, L. Phler, U. Huniar, K. Goss, and A. Klamt, “Prediction of Phospholipid-Water Partition Coefficients of Ionic Organic Chemicals Using the Mechanistic Model COSMOmic,” Journal of Physical Chemistry B **118**, 14833 – 14842 (2014).

[9] G. M. Torrie and J. P. Valleau, “Monte Carlo free energy estimates using non-Boltzmann sampling: Application to the sub-critical Lennard-Jones fluid.” Chemical Physical Letters **28**, 578–581 (1974).

[10] J. Shankar, J. M. Rosenberg, D. Bouzida, R. H. Swendson, and P. A. Kollman, “The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method.” Journal of Computational Chemistry **13**, 1011–1021 (1992).

[11] J. Kastner, “Umbrella Sampling,” WIREs Comput Mol Sci **1**, 932–942 (2011).

[12] C. Jarzynski, “Nonequilibrium equality for free energy differences,” Physical Review Letters **78**, 2690–2693 (1997).

[13] S. Park and K. Schulten, “Calculating potentials of mean force from steered molecular dynamics simulations,” Journal of Chemical Physics **120**, 5946 (2004).

[14] D. Collin, F. Ritort, C. Jarzynski, S. B. Smith, I. T. Jr, and C. Busamante, “Nonequilibrium measurements of free energy differences for microscopically reversible markovian systems,” Nature Letters **90** (1998).

[15] T. Bastug, P. Chen, S. M. Patra, and S. Kuyucak, “Potential of mean force calculations of ligand binding to ion channels from Jarzynski’s equality and umbrella sampling,” Journal of Chemical Physics **128**, 155104 (2008).

[16] T. Bastug and S. Kuyucak, “Application of Jarzynski’s equality in simple versus complex systems,” Chemical Physics Letters **436**, 383 – 387 (2007).

[17] M. Orsi and J. W. Essex, “Physical properties of mixed bilayers containing lamellar and nonlamellar lipids: insights from coarse-grain molecular dynamics simulations,” Faraday Discussions , 249 – 272.

[18] J. Wang, R. M. Wang, J. W. Cadwell, P. A. Kollman, and D. A. Case, “Development and testing of a general AMBER force field,” J. Comput. Chem **25**, 1157–1174 (2004).

[19] W. Humphrey, A. Dalke, and K. Schulten, “VMD – Visual Molecular Dynamics,” Journal of Molecular Graphics **14**, 33 – 38 (1996).

[20] S. Plimpton, “Fast Parallel Algorithms for Short-Range Molecular Dynamics,” Journal of Computational Physics **117**, 1–19 (1995).

[21] M. Tuckerman, B. J. Berne, and G. J. Martyna, "Reversible multiple time scale molecular dynamics," Journal of Chemical Physics **97**, 1990 (1992).

[22] "WHAM: the weighted histogram analysis method," `http://membrane.urmc.rochester.edu/content/wham`.

[23] J. Gore, F. Ritort, and C. Busamante, "Bias and error in estimates of equilibrium free-energy differences from nonequilibrium measurements," Proceedings of the National Academy of Sciences of the United States of America **100**, 12564 – 12569 (2003).

[24] X. Daura, R. Affentranger, and A. E. Mark, "On the Relative Merits of Equilibrium and Non-Equilibrium Simulations for the Estimation of Free-Energy Differences," ChemPhysChem **11**, 3734 – 3743 (2010).

[25] T. V. Pogorelov, J. V. Vermaas, M. J. Arcario, and E. Tajkhorshid, "Partitioning of Amino Acids into Model Membrane: Capturing the Interface," The Journal of Physical Chemistry B **118**, 1481–1492 (2014).

[26] J. L. MacCallum, W. F. D. Bennett, and D. P. Tieleman, Biophysical Journal **94**, 3393 – 3404 (2008).

[27] L. Monticelli, S. K. Kandasamy, X. Periole, R. G. Larson, D. P. Tieleman, and S. J. Marrink, "The MARTINI Coarse-Grained Force Field: Extension to Proteins," Journal of Chemical Theory and Computation **4**, 819 – 834 (2008).

[28] A. Warshel, P. K. Sharma, M. Kato, and W. W. Parson, "Modeling electrostatic effects in proteins," Biochimica et Biophysica Acta **1764**, 1647 – 1676 (2006).

[29] O. Perisic and H. Lu, "On the Improvement of Free-Energy Calculation from Steered Molecular Dynamics Simulations Using Adaptive Stochastic Perturbation Protocols," PLoS ONE **9**, e101810 (2014).

[30] I. Echeverria and L. M. Amzel, "Estimation of Free-Energy Differences from Computed Work Distributions: An Application of Jarzynski's Equality," Journal of Physical Chemistry B **116**, 10986 – 10995 (2012).

[31] F. CÃľlerse, L. Lagardere, ÃĽtienne Derat, and J.-P. Piquemal, "Massively parallel implementation of steered molecular dynamics in tinker-hp: Comparisons of polarizable and non-polarizable simulations of realistic systems," (2019), 10.26434/chemrxiv.7771112.v2.