Hybrid Continuum-Particle Method for Fluctuating Lipid Bilayer Membranes with Diffusing Protein Inclusions

Jon K. Sigurdsson
Department of Mathematics
University of California Santa Barbara.

Frank L.H. Brown
Department of Chemistry
University of California Santa Barbara.

Paul J. Atzberger*
Department of Mathematics,
University of California Santa Barbara.

Abstract: Biological membranes contain many types of embedded proteins whose collective organization and functions depend importantly on the mechanical interplay with the lipid bilayer. We introduce new methods at the level of individual proteins embedded within the bilayer in a manner closely related to the *Immersed Boundary Method* [1, 27]. Our approach accounts for the bidirectional coupling between the membranes and proteins, the elastic mechanics of the bilayer, the hydrodynamic interactions, and the thermal fluctuations. For proteins that induce curvature, we show that the bidirectional membrane-protein coupled dynamics has important consequences for the effective diffusivities of embedded proteins. We further show that collective effects arising from different area fractions and curvatures of the embedded proteins impact significantly membrane mechanics. The proposed modeling approach and computational methods are quite general and could be useful in the investigation of a wide variety of phenomena involving membrane-protein interactions.

* *address:* University of California; Department of Mathematics; Santa Barbara, CA 93106; *e-mail:* atzberg@math.ucsb.edu; *phone:* 805 - 679 - 1330.

I. INTRODUCTION

Biological membranes contain many types of embedded proteins whose collective organization and functions depend importantly on the mechanical interplay with the lipid bilayer. Specific examples include clathrin proteins that induce membrane curvature and the formation of vesicles [19], ion channels and pores gated by voltage and mechanical stresses within the bilayer [6, 9, 14, 28, 34], and anchoring proteins that serve as adhesion sites or junctions with cystoskeletal structures [2, 8, 16, 17, 20]. The abundance of proteins also have important collective effects within biological membranes that significantly alter the membrane mechanics relative to a pure lipid bilayer [10, 18, 28, 33]. An important challenge in cell biology is to understand the physical basis by which membrane embedded proteins perform their cellular functions.

To better understand this interplay requires approaches that go beyond a purely continuum mechanics description, but instead incorporate on some level the important discrete roles played by individual proteins and aggregates within membranes. This poses a number of obstacles and challenges. This requires a balance between the level of detail captured, on the one hand, and the the overall tractability of models for mathematical analysis and computational simulation on the other. Additionally, biological membranes are in a solvated fluid environment and have a bending elasticity permitting significant deformations on energy scales comperable to thermal energy requiring that both hydrodynamics and thermal fluctuations be taken into account.

In this work we present a membrane-protein model based on an approach closely related to the *Immersed Boundary Method* [1, 27]. We formulate this mechanical description in Section II. We then perform an analysis of our approach and some recent related models with respect to the principles of statistical mechanics. When introducing the dynamics for the membrane-protein system, some issues related to how the geometry and the coupling are approximated require careful consideration to ensure consistency with statistical mechanics. We discuss this in detail in Section II B. We then show that some further issues can arise from the breaking of translational symmetry when discretizing such models on a lattice for computational simulations. We show this can result in significant artefacts in the protein dynamics that depend on their relative locations to the underlying discretization mesh. We present computational methods to address these issues in Section III. We show that the level of translational symmetry in simulations can be greatly improved by the choice of the coupling kernel function in Section III A. We then introduce a procedure, we refer to as *Phase-Factor Averaging*, to improve further the translation invariance of the numerical methods.

Applications of our model and numerical methods are presented in Section IV. First, we study the effective diffusivities of curvature inducing proteins and make comparisons with the prior results [23, 31] in Section IV A. We then study collective effects for multiple proteins embedded within the bilayer. We present results for the effective diffusivity and radial distribution functions in Section IV B, Section IV C, and Section IV D. For heterogeneous bilayer membranes with embedded proteins, we investigate the bilayer bending elasticity and we study the effective bending modulus as the area fraction of the embedded proteins is varied in Section IV E. Overall, the proposed modeling approach and computational methods are quite general and we expect they could be useful in the investigation of a wide variety of phenomena involving membrane-protein interactions.

II. MODEL FOR MEMBRANES WITH EMBEDDED PROTEINS

We model a membrane with embedded proteins through a hybrid approach using a particle description for the individual proteins and an augmented continuum mechanics description of the lipid bilayer. We consider in this work a nearly flat membrane that has a configuration that can be described in the Monge-Gauge by the height function $h(x, y)$ parameterized over the xy-coordinate plane. The diffusing proteins are embedded within the membrane and can be described by the constrained positions $\mathbf{s} = (x, y, h(x, y))$.

A. Hamiltonian for Membrane-Protein Mechanics

A common approach to account for the mechanics of a pure lipid bilayer membrane with near vanishing tension is the Helfrich free energy [15]

$$\mathcal{H}_0[h] = \frac{1}{2}\int_{\mathcal{A}_\perp} d\mathbf{x}\left[K_m\left(\mathcal{C}_m[h]\right)^2 + 2K'_m\mathcal{G}_m[h]\right]. \tag{1}$$

We approximate the membrane to leading order relative to a flat sheet and the curvature contributions to second order in $h(x, y)$. The $\mathcal{A}_\perp$ denotes the projected area of the membrane to the xy-coordinate plane and $\mathbf{x} = (x, y)$. The mean curvature is given by $\mathcal{C}_m[h](\mathbf{x}) = \nabla^2 h(\mathbf{x})$ with modulus K_m. The Gaussian curvature is given by $\mathcal{G}_m[h](\mathbf{x}) = \partial_{xx}h(\mathbf{x})\partial_{yy}h(\mathbf{x}) - \partial_{xy}h(\mathbf{x})\partial_{yx}h(\mathbf{x})$ with modulus K'_m.

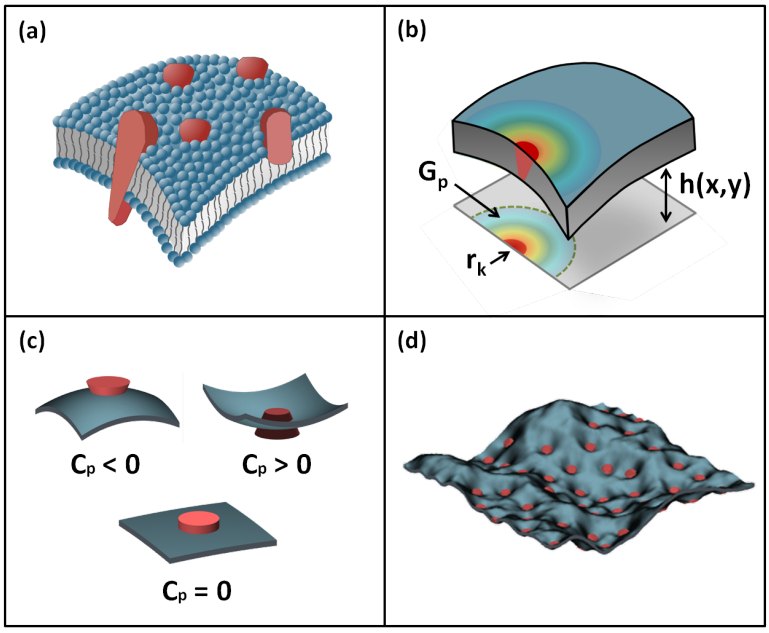

FIG. 1: (a) Biological membranes are a complex mixture of lipids and proteins and other small molecules. Proteins may significantly influence the local bilayer structure and membrane mechanics [33]. (b) To account for how proteins may augment locally the mechanical properties of the bilayer, we introduce in the Hamiltonian a kernel function G_p extending over a range of influence surrounding a protein. The configuration of the membrane is represented as a height function $h(x, y)$ and the protein location projected to the xy-plane by r_k for the k^{th} protein. (c) Proteins may induce a number of different types of local curvature in the membrane. Shown are the cases associated with proteins that augment the local mean curvature C_p. (d) Many invidual and collective mechanical effects of proteins that act on the membrane can be captured in a tractable manner using this approach. This includes membrane mediated protein interactions and forces associated with protein induced deformations of the bilayer.

Within the membrane, the presence of a protein augments through its interactions with the surrounding lipids the local material properties of the bilayer [26, 28], see Figure 1(a). This can make significant contributions to the local bending elasticity and locally preferred curvatures [3, 28, 31, 33]. We account for these effects by introducing the interaction energy

$$\mathcal{H}_{\rm int}[h, \mathbf{r}_k] = \frac{1}{2}\int_{\mathcal{A}_\perp} d\mathbf{x} G_p(\mathbf{x} - \mathbf{r}_k)\Big(K_p(\mathcal{C}_m[h] - 2\mathcal{C}_p)^2 - K_m(\mathcal{C}_m[h])^2 + 2(K'_p - K'_m)\mathcal{G}_m[h]\Big). \tag{2}$$

The configuration of all of the proteins is described in terms of the composite vector $\mathbf{r}$, where $[\mathbf{r}]_k = \mathbf{r}_k = (x_k, y_k)$ gives the location of the k^{th} protein when projected to the xy-plane, so $\mathbf{s}_k = (\mathbf{r}_k, h(\mathbf{r}_k))$. The $\mathcal{C}_p$ denotes the spontaneous mean curvature induced locally by a protein, the factor 2 in front of $\mathcal{C}_p$ represents the two dimensions of $\mathcal{A}_\perp$ and hence $1/\mathcal{C}_p$ the radius of curvature. The elastic moduli locally augmented by the protein is denoted by K_p for the mean curvature and K'_p for the Gaussian curvature. The kernel function $G_p(\mathbf{x})$ accounts for the local range of influence over which the protein augments the bilayer (as tracked by the protein marker), see Figure 1(b). This kernel function is chosen so that it integrates to the area A_p and vanishes outside of a small neighborhood of the protein location. Our specific choice for G_p will be discussed in more detail in Section III A.

The total energy of a configuration of the hybrid membrane-protein system is

$$\mathcal{H}[h, \mathbf{r}] = \mathcal{H}_0[h] + \sum_{k=1}^{N} \mathcal{H}_{\rm int}[h, \mathbf{r}_k] + \mathcal{H}_{\rm pp}[\mathbf{r}]. \tag{3}$$

We have added the contribution to the energy $\mathcal{H}_{\rm pp}[\mathbf{r}]$ to account for direct protein-protein interactions. This term can be motivated by considering the possible contributions arising from sterics, electrostatics, or internal protein mechanics. Here we shall primarily use this term to account for sterics and we give a specific form in Section IV B.

In summary, the Hamiltonian given by equation 3 provides a description of the mechanics that takes into account the bending elasticity of the membrane, the local influence of proteins on augmenting the bilayer, and the protein-protein interactions. This captures salient features of the mechanics of how proteins interact and collectively augment

the lipid bilayer mechanics, see Figure 1(d). We shall discuss this model in more detail in the subsequent sections. Typical values used for the physical parameters appearing in equations 1– 3 can be found in Table I.

B. Membrane-Protein Dynamics

Biological membranes are solvated in a fluid environment and of interest at temperatures where the thermal energy is comparable to the bending energy [28]. This requires accounting in the dynamics for the effects of both hydrodynamic coupling and thermal fluctuations. The time-scales associated with the momentum relaxation of the fluid and membrane on the length-scale of the protein-protein interactions is rapid relative to the time-scale for changes in the deformation of the membrane and changes in the collective configuration of the proteins. In this regime, the viscous stresses of the solvent fluid and membrane balance rapidly with exerted forces making a treatment using overdamped dynamics appropriate [12].

Some recent models have been proposed for membrane-protein dynamics in the overdamped regime [23, 31]. Two central issues arise. The first concerns how the models account for the simultaneous motions of both the membrane and protein that can result in dynamics that are not consistent with the Gibbs-Boltzmann ensemble. The second issue concerns the Hamiltonians used to describe the mechanics in the periodic Monge-Gauge description. The Hamiltonians are truncated to leading order in the height function, remove geometric metric factors, or neglect potentially important contributions arising from the presence of the proteins [23, 31]. Depending on how such approximations are treated, models can be inconsistent with the principles of statistical mechanics. To better understand these issues, we perform analysis of the overdamped dynamical equations used to account for the simultaneous membrane-protein motions. This analysis provides general guidelines and specific conditions useful in the formulation of such models. We then use these conditions to obtain a specific model for the membrane-protein dynamics.

1. *Statistical Mechanics and Membrane-Protein Dynamics*

For the purpose of investigating dynamics, it is convenient to express the membrane-protein stochastic dynamics as

$$\frac{\partial \mathbf{Z}}{\partial t} = M(-\nabla_Z \mathcal{H}) + \mathbf{w} + \mathbf{g}. \tag{4}$$

$$M = \begin{bmatrix} M_{hh} & M_{hr} \\ M_{rh} & M_{rr} \end{bmatrix}. \tag{5}$$

In this notation $\mathbf{Z} = [h, \mathbf{r}]$ denotes the composite vector of all degrees of freedom of the membrane-protein system, namely, the membrane configuration h and all protein positions $\mathbf{r}$. The $\nabla_Z \mathcal{H} = [\delta\mathcal{H}/\delta h, \nabla_{\mathbf{r}}\mathcal{H}]$ denotes the composite vector of all derivatives of the Hamiltonian. The M denotes the grand mobility tensor giving the force response of the membrane and proteins. The $\mathbf{g}$ denotes the stochastic terms accounting for thermal fluctuations. We shall assume throughout that $\mathbf{g}$ is Gaussian with mean 0 and covariance $G = \langle \mathbf{g}\mathbf{g}^T \rangle$. Throughout, all Stochastic Differential Equations (SDEs) should be interpreted in the sense of the Ito Calculus [12, 25]. The $\mathbf{w}$ drift term arises from possible variability of the strength of the thermal fluctuations that depend on the state of the system and the reduction of the inertial dynamics to the overdamped regime [12, 35]. The Gibbs-Boltzmann distribution is required to be invariant under these dynamics with detailed balance as a consequence of the microscopic reversability of the physical processes involved. This yields the following important relationships between the mobility, thermal fluctuations, and thermal drift

$$G = 2k_B T M \tag{6}$$

$$\mathbf{w} = \frac{1}{2}\nabla \cdot G + \frac{1}{2} G \nabla \log(\sqrt{\mathfrak{g}}). \tag{7}$$

The $\mathfrak{g}$ denotes the metric factor associated with the geometry of the membrane and the projection of the protein location onto the reference plane, see Appendix A. The membrane-protein stochastic dynamics consistent with these principles can then be expressed as

$$\frac{\partial \mathbf{Z}}{\partial t} = M(-\nabla_Z \tilde{\mathcal{H}}) + \frac{1}{2}\nabla \cdot G + \sqrt{2k_B T M}\xi. \tag{8}$$

The $\tilde{\mathcal{H}} = \mathcal{H} - k_B T \log(\sqrt{\mathfrak{g}})$ denotes the effective "free-energy" of states when taking the metric factor into account and ξ is the standard Gaussian random field with mean zero, variance one, and no cross-correlations. This is derived and discussed in more detail for membrane-protein systems in Appendix A.

The conditions of equation 6 and equation 7 provide useful guidance when formulating models of the membrane-protein stochastic dynamics to ensure consistency with statistical mechanics. Unfortunately, some prior models proposed for the membrane-protein stochastic dynamics do not satisfy these conditions. In the models of [23, 31], higher order geometric terms are incorporated in the protein dynamics motivated by diffusion on a static curved surface. The membrane-protein dynamics have the thermally induced drift terms corresponding to $\tilde{\mathbf{w}}_h = 0$ and $\tilde{\mathbf{w}}_r = D_0(\mathbf{g}^{-1} : \nabla^2 h)\nabla h/g$, where the Hessian is $(\nabla^2 h)_{ij} = \partial_i\partial_j h$ and the tensor dot product is $A : B = A_{ij}B_{ij}$. In contrast to a static curved surface, the simultaneous membrane-protein stochastic dynamics of these models violate condition 7. In particular, with the grand mobility tensor corresponding to the models, detailed balance for the Gibbs-Boltzmann distribution of equation A7 would require the membrane dynamics to have the thermal drift term $\mathbf{w}_h(\mathbf{x}) = \frac{1}{2}G_{hh}\cdot(\delta/\delta h)\left(\log(\sqrt{\mathfrak{g}})\right) = (k_BT/\mathfrak{g})\,\nabla_{\mathbf{x}}\Lambda(\mathbf{x}-\mathbf{r})\cdot\nabla_{\mathbf{x}}h(\mathbf{r})$ where $\mathbf{r}$ is the protein location and $\Lambda(\mathbf{x}) = 1/8\pi\eta|\mathbf{x}|$ the Oseen tensor component with η the dynamic fluid viscosity [21]. Given the approximations made in treating the geometry, one may also consider the case when treating the Gibbs-Boltzmann distribution with $\mathfrak{g} \approx 1$. However, even in this case, inconsistent results for these models is obtained, since while $\mathbf{w}_h = 0$, this would now require $\mathbf{w}_r = \frac{1}{2}\nabla\cdot G_{rr}$ which is not the same as the above $\tilde{\mathbf{w}}_r$ used in the models.

These inconsistencies highlight that the approximations for treating the geometric contributions of the membrane are inconsistent in different parts of these models. For instance, the membrane dynamical equations only treats the leading order contributions relative to a flat-sheet, while the protein dynamics treats the full contributions of the geometry of the curved membrane surface. In summary, our analysis shows the resulting mobility tensor and drift terms of the models [23, 31] yield collective membrane-protein dynamics inconsistent with detailed balance. To ensure consistency, the conditions of equation 6 and 7 provide useful guidance in the modeling of the membrane-protein dynamics.

2. *Our Model for the Membrane-Protein Dynamics*

To avoid the issues that have arisen for previous models, we shall treat for the purposes of both the membrane and protein dynamics the membrane geometry only to leading order as a nearly flat sheet $\nabla h \ll 1$. The Gibbs-Boltzmann distribution for this approximation has the form

$$\rho(\mathbf{z}) = \frac{1}{\mathcal{Z}}e^{-\mathcal{H}/k_BT} \tag{9}$$

where we treat $\mathfrak{g} \approx 1$ (see Appendix A). We treat in a manner similar to prior models the hydrodynamic coupling of the membrane and the protein drag friction [23, 31]. We consider the grand mobility tensor M with components of the form $M_{hh} = \Lambda*$ (the $*$ denotes convolution), $M_{hr} = M_{rh} = 0$, and $M_{rr} = D_0/k_BT$, as in equation 5. By employing the conditions 6 and 7, we obtain the membrane dynamics

$$\dot{h}(\mathbf{x},t) = \int \Lambda(\mathbf{y}-\mathbf{x})F(\mathbf{y},t)d\mathbf{y} + \sqrt{2k_BT\Lambda}\xi(\mathbf{x},t). \tag{10}$$

The membrane force density is denoted by $F(\mathbf{x},t) = -\left(\delta\mathcal{H}/\delta h\right)(\mathbf{x},t)$. The Λ operator denotes the hydrodynamic coupling of the membrane fluid-elastic sheet. Here, the $\Lambda*$ denotes the convolution with the Oseen tensor $\Lambda(\mathbf{x}) = 1/8\pi\eta|\mathbf{x}|$. This gives the hydrodynamic coupling given by the normal component of the Oseen tensor in the regime of small deformations of a flat-sheet. The η denotes the dynamic fluid viscosity [21]. The last term in equation 10 accounts for thermal fluctuations in a manner consistent with the fluctuation-dissipation principle, see condition 6. This gives the covariance $2k_BTM$ for the thermal forces and that there are no thermally induced drifts, $\mathbf{w}_h = 0$, $\mathbf{w}_r = 0$. The $\xi(\mathbf{x},t)$ denotes a Gaussian white noise process with $\langle\xi(\mathbf{x},t)\rangle = 0$ and $\langle\xi(\mathbf{x},t)\xi(\mathbf{x}',t')\rangle = \delta(\mathbf{x}-\mathbf{x}')\delta(t-t')$.

We remark that the form of the membrane hydrodynamic response treats the normal components as decoupled from the transverse components. Our justification for this is similar to the one given in [22]. When treating the membrane hydrodynamic response to linear order, the normal and transverse modes are decoupled. In particular, the hydrodynamic flows generated by the bending modes of the membrane exert to linear order shear stresses σ that have zero transverse components $\sigma_{xz} = \sigma_{yz} = 0$. As a consequence, only the z-component $\sigma_{zz} = 4\pi\eta|q|\dot{h}_q$ of the hydrodynamic stress plays a role in balancing the membrane force density f, $\sigma_{zz} = 4\pi\eta|q|\dot{h}_q = f_q$. This determines the membrane dynamics and the form of the hydrodynamic mobility tensor, $\dot{h}_q = f_q/(4\pi\eta|q|)$. In real-space, the equations of motion become equation 10. A more detailed derivation of the hydrodynamic in-plane and bending responses can be found in [22].

For the protein dynamics, we have

$$\dot{\mathbf{r}}_k = \frac{D_0}{k_BT}\mathbf{f}_k + \sqrt{2D_0}\boldsymbol{\eta}(t). \tag{11}$$

The D_0 denotes the bare protein diffusion constant. The force acting on the protein is given by $\mathbf{f}_k = -\nabla_{\mathbf{r}_k}\mathcal{H}$. The last term accounts for thermal fluctuations in a manner consistent with the fluctuation-dissipation principle, see condition 6. The $\boldsymbol{\eta}(t)$ denotes a Gaussian white noise process with $\langle\boldsymbol{\eta}(t)\rangle = 0$ and $\langle\boldsymbol{\eta}(t)\boldsymbol{\eta}^*(t')\rangle = \mathcal{I}\delta(t-t')$, where the superscript $*$ denotes vector transpose and $\mathcal{I}$ denotes the identity matrix.

Our specific membrane-protein dynamics of equation 10 and 11 were derived to satisfy conditions 6 and 7. As a consequence, our membrane-protein stochastic dynamics are ensured to have the Gibbs-Boltzmann distribution invariant with detailed balance. This ensures the fluctuations yield results consistent with equilibrium statistical mechanics.

III. COMPUTATIONAL METHODS

To perform computational simulations the spatial and temporal features of the membrane-protein system are discretized. The membrane configuration is represented by the height function $h(x, y)$. We discretize this by using the Fourier basis and approximating the configuration spectrally using a finite number of modes

$$h(\mathbf{x}) = \frac{1}{L^2}\sum_{\mathbf{q}} h_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{x}} \tag{12}$$

where $\mathbf{q} = (q_1, q_2)$, $q^{(j)} = 2\pi j/L$ and $j = -(M-1)/2, \ldots, (M-1)/2$. The modes are represented through the values the membrane height function attains on a regular lattice of sites of the periodic domain, see Figure 2(a). This allows for $h_{\mathbf{q}}$ to be obtained efficiently using a discrete Fast Fourier Transform [7, 29]. Since the height function is real-valued, the coefficients $h_{\mathbf{q}}$ are constrained with $h_{\mathbf{q}} = \bar{h}_{-\mathbf{q}}$. For the degrees of freedom of the membrane in Fourier space, we adopt the convention of using only a subset of the modes and setting the others using this conjugacy condition, see Figure 2(b).

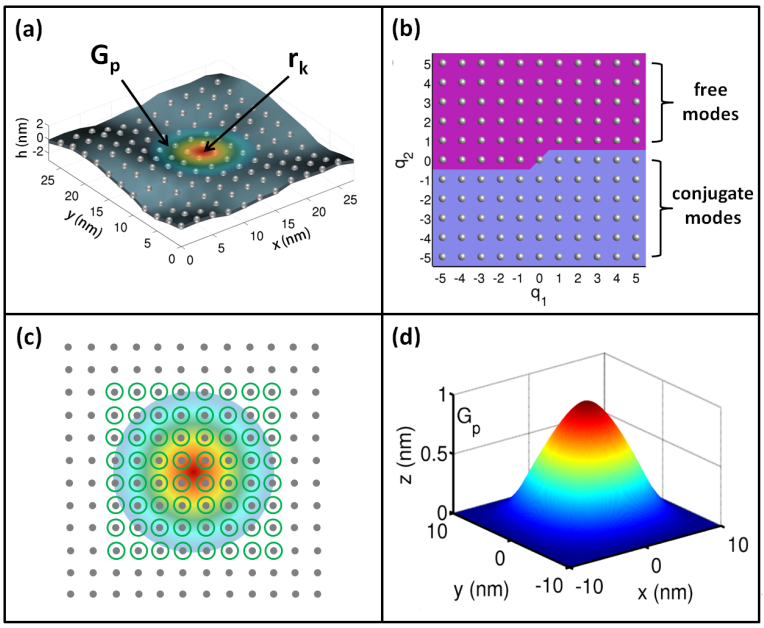

FIG. 2: (a) The configuration of the membrane is represented using a discrete Fourier basis with M modes in each spatial direction. This is equivalently represented in real-space by the values of the height function $h(x, y)$ sampled on a regular lattice. The protein location is represented by projection to the xy-plane by $\mathbf{r}_k$ for the k^{th} protein. The kernel function G_p represents the range of local influence of the protein on the bilayer. (b) Since the Fourier series must represent a real-valued height function only a subset of the modes are "free" with the others constrained to be "conjugate" to the free modes. Shown is the convention used in computations for the free and conjugate modes. (c) The protein can be located at off-lattice locations. We use the kernel function G_p to communicate between values at off-lattice protein locations with values at the surrounding lattice sites. (d) The kernel function G_p used in our computations.

The Hamiltonian is discretized by using the trapezoidal quadrature [4]. All forces are obtained from the discretized

Hamiltonian by taking variations with respect to the discrete membrane and protein degrees of freedom. This ensures that the variational structure of the membrane-protein system are preserved so that the discretized energy and force are consistent [13]. The particular energy and forces associated with the discretized system will depend on the specific choice made for the kernel function $G_p(\mathbf{r})$. The choice of the kernal function requires some care to avoid the introduction of inadvertent artifacts arising from the lattice discretization that breaks translational symmetry, see Figure 2(c).

To discretize in time, we approximate the membrane stochastic dynamics using the Euler-Maruyama Method in Fourier space

$$h_{\mathbf{q}}^{n+1} = h_{\mathbf{q}}^{n} + \Delta t \Lambda_{\mathbf{q}} F_{\mathbf{q}}^{n} + \sqrt{\Delta t k_B T L^2 \Lambda_{\mathbf{q}}} \xi_{\mathbf{q}}^{n} \tag{13}$$

where $\Delta t = t_{n+1} - t_n$ is the time step. The $F_{\mathbf{q}}^n$ denotes the Fourier representation of the membrane force density at time $t_n = n\Delta t$ of $F_{\mathbf{m}}^n = -\left(\delta \mathcal{H}/\delta h\right)(\mathbf{x}_m, t_n)$, see further discussion in Appendix B. The $\Lambda_{\mathbf{q}} = 1/4\eta q$ denotes the Fourier representation of the hydrodynamic coupling given by the Oseen tensor component $\Lambda(\mathbf{x}) = 1/8\pi\eta|\mathbf{x}|$ with η the dynamic fluid viscosity [21]. The last term accounts for the contributions of the thermal fluctuations over the time step. The $\xi_{\mathbf{q}}^n$ denotes a Gaussian random variable with $\langle \xi_{\mathbf{q}}^n \rangle = 0$ and $\langle \xi_{\mathbf{q}}^n \xi_{\mathbf{q}'}^{n'} \rangle = 2\delta_{\mathbf{q},-\mathbf{q}'}\delta_{n,n'}$.

The protein dynamics are discretized in time using the Euler-Maruyama Method

$$\mathbf{r}_k^{n+1} = \mathbf{r}_k^n + \Delta t \frac{D_0}{k_B T} \mathbf{f}_k + \sqrt{2D_0 \Delta t}\boldsymbol{\eta}^n. \tag{14}$$

The D_0 denotes the bare protein diffusion constant. The force acting on the protein is given by $\mathbf{f}_k^n = -\nabla_{\mathbf{r}_k}\mathcal{H}(t_n)$, where we have discretized the Hamiltonian using a simple quadrature using summation over the available values of h and its derivatives. The last term accounts for the contributions of the thermal fluctuations over the time step. The $\boldsymbol{\eta}^n$ denotes a Gaussian random variable with $\langle \boldsymbol{\eta}^n \rangle = 0$ and $\langle \boldsymbol{\eta}^n \boldsymbol{\eta}^{n'*} \rangle = \mathcal{I}\delta_{n,n'}$, where the superscript $*$ denotes vector transpose and $\mathcal{I}$ denotes the identity matrix.

A. Approximate Translation Invariance of the Membrane-Protein Coupling : Importance of the Kernel Function

An important consideration when discretizing the Hamiltonian is how to treat the coupling of the protein and membrane so that the system behaves translationally invariant to a good approximation. This depends importantly on the discretization lattice used for $\mathcal{H}_{int}$ and the specific choice made for the kernel function $G_p(\mathbf{r})$. To investigate how these factors influence the level of translation invariance, we consider variations in the membrane surface obtained when minimizing the Hamiltonian while fixing the protein location at different shifts relative to the underlying discretization mesh. For a perfectly invariant system the membrane surface should have modulo translation a shape that remains invariant for all translations of the protein and should have exactly the same energy. The underlying discretization lattice breaks this strict translational symmetry.

For a curved protein with positive Gaussian curvature, the membrane-protein interaction induces a bowl-like shape in the membrane surface that tracks the protein location. Variations in the shape of this bowl-like shape and in the associated energy give a strong indication of the level of translation invariance of the system. Consequently, if these energy variations are large they could play a significant practical role in simulations by contributing artificially to the energy landscape experienced by a diffusing protein. In practice, these energy variations can be used as a measure of the quality of any particular implementation of the numerical methods. Small variations relative to k_BT indicate that for the purpose of thermal simulations the numerical methods approximate well the translation invariance of the system. Whereas, large variations indicate that the numerical methods significantly break translation invariance and will suffer rather severe lattice artifacts in simulations.

To design numerical methods that approximate well the translation invariance of the system, there are two natural strategies to consider: (i) attempt to develop a protein kernel function G_p that varies smoothly in the protein location and that has a shape that complements well the membrane-protein coupling to minimize variations. (ii) refine the lattice mesh locally so that it more accurately discretizes the Hamiltonian and offers more resolution to approximate the local membrane shape (which in the continuum limit becomes translationally invariant). To investigate the first approach, we compare three different kernel functions: (i) linear hat kernel function G_{hat}, (ii) truncated Guassian kernel function $G_{gaussian}$, and (iii) Peskin δ-function G_{peskin}, the specific functional forms are given in Appendix C and [1, 27]. To investigate the second approach, we fix the form of the kernel function and refine the discretization lattice. This has the result of increasing the support of the kernel function M_{supp}, which is the number of lattice sites for which the kernel function is non-zero (in each spatial direction). While each of these strategies can be shown to provide useful ways to improve the translation invariance, each has associated computational costs that are important to balance in a practical simulation.

The computational cost associated with refinement of the mesh arises primarily from two sources (i) the increased cost of the FFTs, and (ii) the introduction of additional time-scales into the temporal dynamics limiting the largest time-step Δt that can be taken during simulations. This latter issue arises from the largest frequency Fourier modes of the membrane that relax on a time-scale q^3 in the wavenumber. When refining the lattice spacing by a factor α^{-1}, the time-step Δt is restricted by $\alpha^{-3}\Delta t_0$, where Δt_0 was the bound before refinement. For the FFTs the mesh has N^2 points and after refinement $\alpha^2 N^2$ points which has a cost each time-step that scales like $\alpha^2 N^2 \ln(\alpha^2 N^2)$. As a consequence, the total cost associated with refining the mesh for a simulation performed over a duration of time $[0, T_0]$ scales like $\alpha^5 N^2 \ln(\alpha^2 N^2) T_0$. The factor of α^5 presents a rather severe constraint on how much the mesh can be refined in order to ameliorate the issues associated with the translation invariance. Even a modest refinement by a factor $\alpha^{-1} = 1/2$ increases the simulation cost by $\alpha^5 = 32$.

To benefit from the better translation invariance offered by a more refined mesh, but to achieve a much lower computational cost than outright mesh refinement, we introduce an alternative method. In particular, we perform an averaging procedure to help make more translationally symmetric the numerical calculations performed using the discretization mesh. An important feature of our numerical discretization of the membrane-protein system is that the lattice artifacts that disrupt translation invariance enter primarily when the Fourier representation of the force density of the protein acting on the membrane is computed through the use of the Fast Fourier Transform. This arises since the protein force density is not band-limited and is sampled only on a discrete lattice. As a consequence, the force density is not exactly represented by the finite number of Fourier modes and its discretization depends on the lattice, see equations 2, 13 and 14. To reduce artifacts that arise as this force density is shifted relative to the lattice, we introduce a procedure that samples the force density on several different discrete lattices and uses the phase-factor to perform an averaging to estimate the Fourier coefficient. We refer to this method as *Phase-Factor Averaging*.

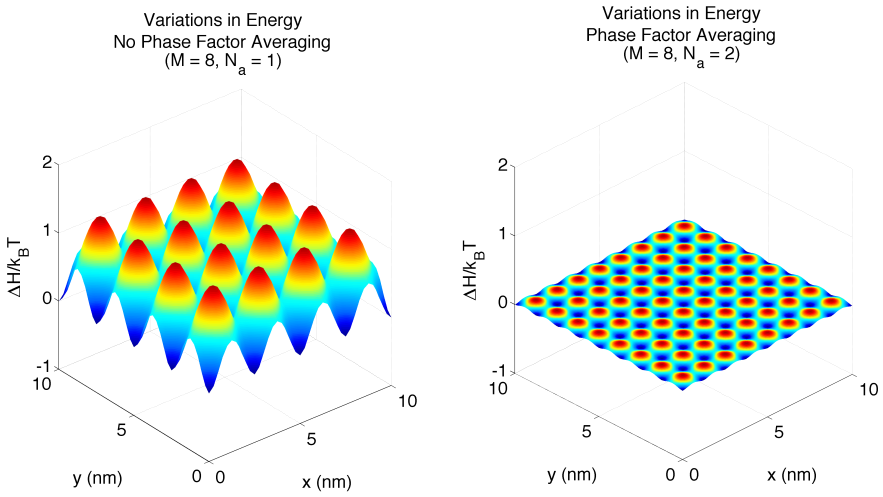

FIG. 3: Issues with Translation Invariance of the Membrane-Protein Coupling. Shown is the variations of the minimum energy of the membrane surface as the protein is shifted relative to the underlying discretization lattice, $\mathcal{C}_p = 0.15\text{nm}^{-1}$. The ΔH is the difference between the energy value at a shifted site relative to being located at a lattice site. When using only the kernel function values on the discretization lattice the energy variations are found to be on the order of $k_B T$ (left panel). When using the phase-factor averaging method to sample values off-lattice, the energy variations are nearly eliminated with variations no more than $0.08 k_B T$ (right panel).

1. *Phase Factor Averaging*

Numerical errors are introduced with respect to the translation invariance of the model in only two parts of the algorithm. This occurs when computing the force density arising from the protein acting on the membrane and when computing the force acting on the protein from the membrane. The issue has to do with the non-band limited form of the resulting protein associated force field which can only be captured approximately by a finite number of Fourier modes, see equations 2, 13 and 14. To reduce artifacts introduced by the discretization mesh used for sampling during the computation of the the discrete Fast Fourier transform, we perform a phase-factor weighted averaging over several shifts of the mesh.

For a general field, the Fourier transform on a periodic domain is given by

$$u_{\mathbf{q}} = [\mathcal{F}u]_{\mathbf{q}} = \int u(\mathbf{x}) e^{-i\mathbf{q}\cdot\mathbf{x}} d\mathbf{x}. \tag{15}$$

Let the shift operator be defined by $\mathcal{T}_{\mathbf{s}} u = u(\mathbf{x}-\mathbf{s})$. The shift operator then introduces a phase factor into the Fourier transform that satisfies

$$[\mathcal{F}u]_{\mathbf{q}} = e^{i\mathbf{q}\cdot\mathbf{s}} [\mathcal{F}(\mathcal{T}_{\mathbf{s}} u)]_{\mathbf{q}}. \tag{16}$$

The full Fourier transform can be expressed exactly in terms of a phase-factor weighted average over any collection of shifts s_j as

$$[\mathcal{F}u]_{\mathbf{q}} = \frac{1}{N_s} \sum_{j=1}^{N_s} e^{i\mathbf{q}\cdot\mathbf{s}_j} [\mathcal{F}(\mathcal{T}_{\mathbf{s}_j} u)]_{\mathbf{q}}. \tag{17}$$

This identity can be usefully employed to achieve efficient numerical methods that better approximate translation invariance. In particular, we replace the exact Fourier transform $\mathcal{F}$ by the discrete Fourier transform $\bar{\mathcal{F}}$ and we compute the Fourier coefficients $\hat{u}_{\mathbf{q}}$ of a general field u as

$$\hat{u}_{\mathbf{q}} = [\bar{\mathcal{F}}u]_{\mathbf{q}} = \frac{1}{N_s} \sum_{j=1}^{N_s} e^{i\mathbf{q}\cdot\mathbf{s}_j} [\tilde{\mathcal{F}}(\mathcal{T}_{\mathbf{s}_j} u)]_{\mathbf{q}}. \tag{18}$$

To obtain the Fourier coefficients of the protein force density field that acts on the membrane we use the phase-factor weighted Fourier transform $\mathbf{F}_{\mathbf{q}} = [\bar{\mathcal{F}}\mathbf{F}]_{\mathbf{q}}$. To handle the force exerted by the membrane acting on a protein $\mathbf{f}_j = d\mathcal{H}/dr_j$, we compute the approximations to the integrals using a quadrature that makes use of the additional values of the height field and its derivatives at all of the shifted lattice locations used in the phase-factor weighted Fourier transform. This modification ensures a higher level of translational symmetry of the computation which now depends only on the relative displacement from the nearest shifted lattice location (as opposed to the original discretization mesh).

In practice, we perform the averaging procedures by introducing N_a equally spaced shifts in each direction. This should be compared with the cost of refining the discretization mesh N_a times, which during each time step costs a factor of N_a^2 for the FFTs and from the introduction of fast temporal dynamics for the highest frequency Fourier modes costs a factor N_a^3 for a total cost of N_a^5, see Section III A. In contrast, the phase-factor averaging procedure only costs N_a^2.

2. *Results for Translation Invariance of Numerical Methods : Comparison for Different Kernel Functions, Protein Curvatures, and Level of Phase-Factor Averaging*

We now discuss how the various kernel functions G_p perform when refining the mesh and when performing phase-factor averaging for proteins of different curvatures. We compare the following (i) linear hat keneral function G_{hat}, (ii) truncated Gaussian kernel function $G_{gaussian}$, and (iii) Peskin kernel function G_{peskin}, see Appendix C and [1, 27]. To investigate the role of mesh refinement, we fix the protein curvature and the functional form of the kernel function and then refine the underlying discretization mesh. To characterize the level of refinement we use the support of the kernel function M_{supp}, which is the number of lattice sites for which the kernel is non-zero in each spatial direction. To investigate how phase-factor averaging impacts the translation invariance, we again study the same cases but estimate the Fourier coefficients using $N_s = N_a^2$ lattice shifts of the mesh. We perform N_a lattice shifts by the amounts $(1/N_a)\Delta x$ in each spatial direction where Δx is the lattice spacing. A useful indicator of the number of function evaluations of the kernel function is $N_* = N_a M_{supp}$. The number of modes used is $M = 4M_{supp} + 1$ and hence the box size $L = M\Delta x$ and the other default parameters are given in Table I. The results of our studies are given in the Table II, Table III, and Table IV.

We find that the choice of kernel function plays a significant role in the level of translation invariance achieved. The hat function is found to provide rather poor performance for the cases considered even when performing a significant amount of phase-factor averaging, see Table II. For example, the case that might be most natural to use in simulations with the support $M_{supp} = 4$ and no phase-factor averaging $N_a = 1$ yields energy variations of $4k_BT$ when the curvature is $C_p = 0.05\text{nm}^{-1}$ and $40k_BT$ when the curvature is $C_p = 0.15\text{nm}^{-1}$. Even when increasing significantly the support to $M_{supp} = 24$ or performing a large number of phase-factor averaging shifts $N_a = 12$ ($N_* = 24$), energy variations persist on the order of k_BT for $C_p = 0.15\text{nm}^{-1}$, see Table II. These significant lattice artifacts appear to arise from

Parameter	Box dimension	Number of proteins	Protein area	Temperature	Bare protein diffusion coefficient	Bilayer bending modulus	Protein bending modulus	Saddle-splay moduli	Protein spontaneous curvature	Solvent viscosity
Symbol	L	N	A_p	T	D_0	K_m	K_p	$K'_{m(p)}$	C_p	η
Value	247.5 nm	0-100	100 nm^2	300 K	0.025 $\frac{\text{nm}^2}{\text{ns}}$	$5k_BT$	$40k_BT$	$-K_{m(p)}$	0 nm^{-1}	$10^{-3}\text{Pa}\cdot\text{s}$

TABLE I: Description of the parameters and default values used in the simulations.

the non-smooth piece-wise linear variation of the weights as the protein shifts relative to the discretization lattice. This results in sensitivity of the membrane shape that changes significantly as the protein shifts.

The truncated Gaussian provides a kernel function with a much smoother behavior for the weights as the protein shifts. For the case using the smallest support and no phase-factor averaging, we obtain $0.4k_BT$ when the curvature is $C_p = 0.05\text{nm}^{-1}$ and $3k_BT$ when the curvature is $C_p = 0.15\text{nm}^{-1}$. In comparison to the hat kernel this already is a dramatic improvement. When increasing the support to $M_{supp} = 24$ or performing phase-factor averaging shifts $N_a = 12$ ($N_* = 24$), energy variations can be brought down to the order of $0.3k_BT$ for $C_p = 0.15\text{nm}^{-1}$, see Table III.

To further improve the translation invariance of the numerical discretization, we introduced the use of the Peskin kernel function [23], which was designed in the context of hydrodynamic calculations to have favorable translation invariance properties [1, 27]. For the case using the smallest support and no phase-factor averaging, we obtain $1.5k_BT$ when the curvature is $C_p = 0.05\text{nm}^{-1}$ and $13k_BT$ when the curvature is $C_p = 0.15\text{nm}^{-1}$. While in comparison to the hat kernel this is an improvement, in comparison with the Gaussian this appears initially significantly worse. However, once phase-factor averaging is performed the Peskin kernel function performs significantly better than the Gaussian kernel. When increasing the support to $M_{supp} = 24$ or performing phase-factor averaging shifts $N_a = 12$ ($N_* = 24$), energy variations can be brought down to the order of $0.01k_BT$ for $C_p = 0.15\text{nm}^{-1}$, see Table IV. Interestingly, very good results are even obtained for the more modest case when $M_{supp} = 16$ or $N_a = 8$ ($N_* = 16$) with energy variations of $0.06k_BT$ for $C_p = 0.15\text{nm}^{-1}$. In contrast, the Gaussian kernel function shows persistent energy variations of $0.5k_BT$. In practical simulations the Peskin kernel function offers the best performance allowing for a modest support and number of phase factor averaging steps to be used to remove effectively artifacts associated with translation invariance.

To illustrate how the phase-factor averaging improves the simulation results, we consider the case when the membrane energy is minimized for each location of a protein. In this case, a protein would diffuse in an energy landscape corresponding to the energy variations we have reported. In the case of a curved protein with $C_p = 0.15\text{nm}^{-1}$ and kernel function G_{peskin} with $M_{supp} = 8$ and no-phase factor averaging the energy landscape is rough with variations on the order of k_BT. As a consequence the diffusivity of the protein is influenced strongly by the lattice artifacts of the discretization. In contrast, a smaller kernel can be used with $M_{supp} = 4$ with phase-factoring averaging $N_a = 4$ to achieve much smaller energy variations on the order $0.08k_BT$. The detailed features of the energy landscape obtained using these two different numerical methods are shown in Figure 3. In summary, the reported results show that phase-factor averaging with the Peskin kernel function provides a computationally efficient and highly effective method to achieve near translation invariance in simulations.

TABLE II: Hat Kernel : Phase Factor Averaging

Hat Kernel Function G_{hat} : Energy Variations $\Delta\widehat{\mathcal{H}}/k_BT$

N_*	N_a	M_{supp} \ C_p	0.05 nm^{-1}	0.10 nm^{-1}	0.15 nm^{-1}
4	1	4	4.6127	18.4509	41.5145
	2	2	3.7613	15.0450	33.8513
8	1	8	1.3217	5.2866	11.8950
	2	4	1.0501	4.2005	9.4512
	4	2	0.8885	3.5538	7.9961
12	1	12	0.6158	2.4633	5.5425
	2	6	0.5203	2.0814	4.6831
	3	4	0.4584	1.8334	4.1252
	4	3	0.4084	1.6336	3.6757
	6	2	0.3908	1.5632	3.5172
16	1	16	0.3540	1.4158	3.1857
	2	8	0.3040	1.2160	2.7360
	4	4	0.2563	1.0251	2.3064
	8	2	0.2190	0.8761	1.9713
20	1	20	0.2292	0.9166	2.0624
	2	10	0.2026	0.8104	1.8234
	4	5	0.1668	0.6671	1.5010
	5	4	0.1636	0.6542	1.4721
	10	2	0.1399	0.5598	1.2595
24	1	24	0.1602	0.6409	1.4420
	2	12	0.1433	0.5732	1.2898
	3	8	0.1334	0.5338	1.2010
	4	6	0.1273	0.5093	1.1458
	6	4	0.1134	0.4537	1.0207
	8	3	0.1015	0.4058	0.9132
	12	2	0.0971	0.3884	0.8739

Where $N_* = N_a M_{supp}$.

TABLE III: Gaussian Kernel : Phase Factor Averaging

Gaussian Kernel Function $G_{gaussian}$: Energy Difference $\Delta\widehat{\mathcal{H}}/k_BT$

N_*	N_a	M_{supp} \ C_p	0.05 nm^{-1}	0.10 nm^{-1}	0.15 nm^{-1}
4	1	4	0.4053	1.6210	3.6473
	2	2	0.2212	0.8848	1.9907
8	1	8	0.2461	0.9842	2.2145
	2	4	0.2072	0.8287	1.8646
	4	2	0.1226	0.4902	1.1030
12	1	12	0.1477	0.5906	1.3289
	2	6	0.1376	0.5506	1.2388
	3	4	0.1229	0.4915	1.1058
	4	3	0.0964	0.3856	0.8675
	6	2	0.0739	0.2957	0.6652
16	1	16	0.1040	0.4161	0.9361
	2	8	0.0994	0.3977	0.8948
	4	4	0.0867	0.3469	0.7806
	8	2	0.0528	0.2110	0.4749
20	1	20	0.0799	0.3196	0.7191
	2	10	0.0770	0.3082	0.6934
	4	5	0.0715	0.2861	0.6438
	5	4	0.0669	0.2674	0.6017
	10	2	0.0410	0.1639	0.3688
24	1	24	0.0647	0.2589	0.5824
	2	12	0.0627	0.2510	0.5647
	3	8	0.0615	0.2460	0.5535
	4	6	0.0599	0.2398	0.5395
	6	4	0.0543	0.2173	0.4890
	8	3	0.0429	0.1717	0.3864
	12	2	0.0335	0.1339	0.3014

Where $N_* = N_a M_{supp}$.

TABLE IV: Peskin Kernel : Phase Factor Averaging

Peskin Kernel Function G_{peskin}: Energy Variations $\Delta\widehat{\mathcal{H}}/k_BT$

N_*	N_a	M_{supp} \ C_p	0.05 nm^{-1}	0.10 nm^{-1}	0.15 nm^{-1}
4	1	4	1.5328	6.1313	13.7954
	2	2	0.6828	2.7313	6.1453
8	1	8	0.1422	0.5689	1.2800
	2	4	0.1067	0.4269	0.9605
	4	2	0.0667	0.2667	0.6000
12	1	12	0.0263	0.1051	0.2365
	2	6	0.0270	0.1080	0.2430
	3	4	0.0241	0.0963	0.2166
	4	3	0.0171	0.0682	0.1535
	6	2	0.0160	0.0642	0.1444
16	1	16	0.0075	0.0300	0.0676
	2	8	0.0094	0.0374	0.0842
	4	4	0.0082	0.0330	0.0742
	8	2	0.0056	0.0225	0.0506
20	1	20	0.0029	0.0116	0.0261
	2	10	0.0039	0.0157	0.0354
	4	5	0.0033	0.0134	0.0301
	5	4	0.0035	0.0142	0.0319
	10	2	0.0024	0.0098	0.0220
24	1	24	0.0013	0.0054	0.0121
	2	12	0.0018	0.0073	0.0165
	3	8	0.0020	0.0079	0.0177
	4	6	0.0019	0.0077	0.0173
	6	4	0.0018	0.0071	0.0159
	8	3	0.0013	0.0052	0.0117
	12	2	0.0012	0.0049	0.0110

Where $N_* = N_a M_{supp}$.

IV. APPLICATIONS

Biological membranes contain diverse types of embedded proteins ranging from gated channels for ions and small molecules to anchors that interact with the cytoskeleton [2]. How such structures are assembled, regulated, and perform their various functions involve significant mechanical interplay with the lipid bilayer membrane [18, 33]. Additionally, the abundance of such proteins are expected to have important collective effects that significantly alter the mechanical behaviors of the membrane relative to a pure lipid bilayer [10, 18, 28, 33].

We perform a number of simulations to validate the proposed membrane-protein model and to show some of the types of phenomena that may be investigated with our methods. We first study the diffusivity of proteins that induce a spontaneous curvature of the bilayer and investigate the differences that arise when compared to proteins that diffuse on a static surface. We then study how the diffusivity of a protein depends on the area fraction of proteins within the bilayer and how steric interactions with the other proteins hinder the diffusive motion. We then investigate the contributions of the membrane mechanics and fluctuations to the collective interactions between the proteins and the equilibrium ensemble of their arrangements within the bilayer. To study how the mechanics of the bilayer is impacted by the presence of proteins, we investigate the homogenized elasticity of the bilayer through a spectral analysis of the undulation fluctuations of the membrane. These applications give a demonstration of the capabilities of the proposed membrane-protein modeling approach and computational methods.

A. Diffusion of a Protein that Induces Local Curvature

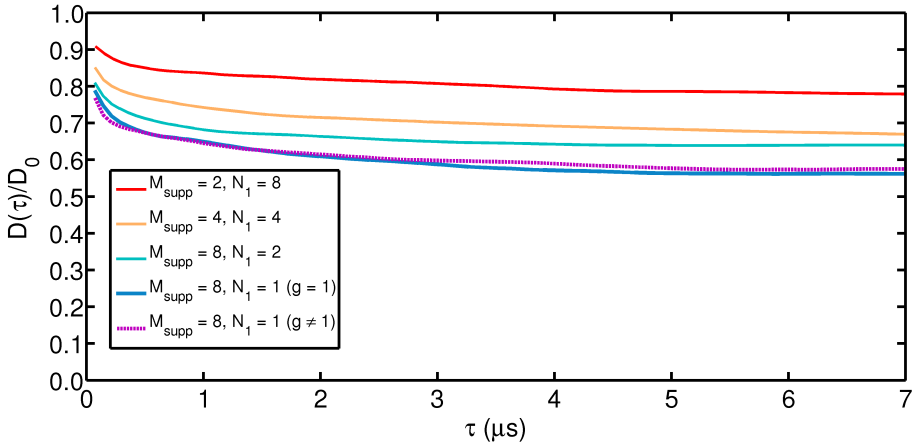

FIG. 4: Diffusion of a Curvature Inducing Protein. The time dependent diffusion coefficient $D(\tau)$ is shown of a protein that induces a local curvature in the bilayer, $\mathcal{C}_p = 0.15\text{nm}^{-1}$, $D_0 = 0.005\text{nm}^2/\text{ns}$. This is normalized by comparing to the bare diffusion coefficient D_0 of a protein diffusing within a static flat surface. The diffusion coefficient obtained for the membrane-protein dynamics given in equation 10-11 is shown when using different supports for the kernel function M_{supp} and number of phase-factor shifts N_a. The membrane-protein coupling is influenced significantly by these choices through the effective induced shape of the membrane shape that tracks the protein location. In the numerical methods, these factors also influence the level of translation invariance achieved in simulations, as discussed in Section III A. The diffusion coefficient obtained for the membrane-protein dynamics given in the prior work [23] corresponds to $(M_{supp} = 8, N_a = 1, (g \neq 1))$. The diffusion using the dynamics that exhibits detailed balance by excluding the geometric term corresponds to $(M_{supp} = 8, N_a = 1, (g = 1))$. An even more substantial difference arises obtained when using a kernel function that better exhibits translation invariance in simulations, see $(M_{supp} = 8, N_a = 2)$. The larger diffusivity arises from the removal of energy barriers and reduction of the energy variations over lattice shifts by a factor of about 15 under the phase-factor averaging, see Table IV.

Many proteins embedded within the bilayer are expected to locally induce a preferred curvature of the bilayer. An interesting question is how the coupling of the protein mechanics to the bilayer mechanics influences the protein diffusivity. This problem has been considered in the prior works [23, 31]. An important feature of the membrane-protein model proposed here is that we take into account the bidirectional mechanical coupling between the membrane and protein. Here and in our prior work [23], we show neglecting this important bidirectionality in the mechanics can result in problems that violate the principles of statistical mechanics, see Section II B 1. In fact, many approximations

are inherent when using the Helfrich Free Energy and Monge-Gauge to capture the leading order behavior of the membrane mechanics. In particular, in some models higher order geometric contributions from the membrane surface are included in the protein dynamics, that while valid for a static membrane surface, turn out to be inconsistent with the other approximations implicit in the model when both the protein and membrane undergo simultaneously stochastic fluctuations. An important feature of the membrane-protein model proposed here is that rigorous statistical mechanical analysis has been done to ensure these approximations enter in the model in a consistent fashion, see Section II B 1. Through our analysis of translation invariance of the numerical methods we have also found that significant energy variations can arise from lattice artefacts that influence diffusivity depending on the choice of kernel and level of mesh refinement. These considerations highlight the importance in considering carefully the underlying numerical methods and statistical mechanics when performing such simulations.

Using these more robust numerical methods, we perform simulations of the diffusive motion of a protein that induce a local curvature on the membrane. We consider the time-dependent diffusion coefficient

$$D(\tau) = \frac{\langle |\mathbf{r}(\tau) - \mathbf{r}(0)|^2 \rangle}{4\tau}, \tag{19}$$

where τ is the delay time and $\mathbf{r}(\tau) - \mathbf{r}(0)$ is the displacement in the protein position. We consider the case when the protein induces a rather strong local curvature on the bilayer $\mathcal{C}_p = 0.15\text{nm}^{-1}$, $D_0 = 0.005\text{nm}^2/\text{ns}$. The other simulations parameters used can be found in Table I and Table V. Our results are shown in Figure 4.

An important finding is that the breaking of translation invariance can introduce lattice dependent energy variations that play a significant role in augmenting the protein diffusive motion. For instance in our prior work [23], the energy variations are found to be on the order of k_BT and correspond to the case $M_{supp} = 8, N_a = 1$ in Figure 4. This prior numerical method effectively presents the diffusing protein with an effective unphysical energy landscape that has barriers on the order of k_BT between preferred minimal energy locations that depend on the lattice used. As a consequence, this artificially augments the protein motion and the overall diffusivity. When using our new numerical methods with the same kernel function as in the prior work [23] but with phase factor averaging shifts ($N_a = 2$), we find the energy variations are greatly reduced and effectively removed. For a comparison of the effective energy landscape for the protein in the cases with and without phase factor averaging performed, see Figure 3. As would be expected, the corrected results using our new numerical methods give a significantly larger protein diffusivity relative to [23], see $M_{supp} = 8, N_a = 2$ in Figure 4. In this case, the difference between the methods concerning the geometric terms play less of a role in influencing the diffusivity than the predominant issues with the energy variations arising from the discretization mesh. The comparison when retaining the unphysical energy variations but only augmenting the geometric term are given by $M_{supp} = 8, N_a = 1, (g \neq 1)$ (same as [23]) and $M_{supp} = 8, N_a = 1, (g = 1)$ (dynamics of Section II B 2) where g denotes whether the metric factor is included $g \neq 1$ or not $g = 1$, see Figure 4. These results show the importance of making an appropriate choice for the kernel and performing the phase-factor averaging to achieve reliable numerical results.

Parameter	Lattice spacing	Number of modes	Time step
Symbol	Δx	$M = L/\Delta x$	Δt
Value	2.5 nm	99	≈ 7 ps

TABLE V: Discretization Parameters

B. Sterics $\mathcal{H}_{\text{pp}}$ for Multiple Proteins

When considering membrane-protein systems with multiple proteins it is important to take into account the role of sterics. This will be done through the contributions of $\mathcal{H}_{\text{pp}}$ to the Hamiltonian in equation 3. For this purpose, we use the pairwise repulsive interactions

$$\mathcal{H}_{\text{pp}}[\mathbf{r}] = \sum_k \sum_{j<k} \phi_{\text{pp}}(|\mathbf{r}_k - \mathbf{r}_j|) \tag{20}$$

where $\phi_{\rm pp}(r) = C/(r-r_1)^5$. For use in practice, such potentials must be regularized [11]. For this purpose, we use a shift and truncation on the interval $[r_1, r_2]$ which yields the potential

$$\phi_{\rm pp}(r) = \int_0^r \min(f_{max}, f(s))ds \tag{21}$$

$$f(r) = \begin{cases} \infty & \text{if } 0 \le r \le r_1, \\ \frac{\alpha k_B T}{(r_2-r_1)^6}\left(\frac{(r_2-r_1)^6}{(r-r_1)^6} - 1\right) & \text{if } r_1 < r < r_2, \\ 0 & \text{if } r_2 \le r. \end{cases} \tag{22}$$

In the presented simulations, we use the parameters given in table VI.

Parameter	Repulsive radius	Inner radius	Cut off radius	Interaction strength	Maximum displacement	Maximum force	Upper bound
Symbol	$R_{\rm p}$	r_1	r_2	α	$r_{\rm m}$	$f_{\rm max}$	$\Delta t \le \Delta t^*$
Value	$4\Delta x$	$\frac{30}{16}R_{\rm p}$	$\frac{50}{16}R_{\rm p}$	200	2 nm	$\frac{k_B T}{D_0}\frac{r_{\rm m}}{\Delta t}$	10 ps

TABLE VI: Parameters for Repulsive Steric Interactions $\mathcal{H}_{\rm pp}$

C. Diffusivity of Multiple Proteins within a Membrane

To investigate the contributions of sterics to the diffusivity of proteins within the bilayer, we consider how the effective single protein diffusivity depends on the area fraction of proteins within the bilayer. To avoid additional contributions arising from the curvature mediated interactions between proteins, we initially study the effective diffusion coefficient for flat proteins, $C_p = 0\text{nm}^{-1}$. We vary the area fraction Ψ over the range 12% to 52%, which corresponds to having the number of proteins N vary between 25 to 100 for the parameters in Table V. To sample the diffusivities in this system, we ran one long simulation over 5×10^7 time steps. We then used the configuration of this simulation at time steps $(1 + 0.5n) \times 10^7$, $n = 0, 1, \ldots, 8$, as the initial condition for nine new simulations that were run in parallel for 5×10^7 time steps. This was done for each area fraction to sample the diffusion of the proteins. Our results are shown in Figure 5.

For the small volume fractions we find that the protein diffusivity is close to the bare diffusivity of the protein. As the volume fraction increases we find that the steric interactions result in local trapping of the protein by neighboring proteins requiring coordinated motions to allow the protein to diffuse over longer distances. Interestingly, we find that the decrease in the diffusivity resulting from sterics follows a linear trend in the area fractions over the range considered. This is shown in Figure 6. We remark that in these results only the sterics and membrane mechanics have been considered and the possible contributions of collective hydrodynamic interactions in the lateral direction directly between the proteins to the diffusivity have been neglected in this initial study.

D. Arrangements of Multiple Proteins within the Bilayer : Radial Distribution Function and Potential of Mean Force

When multiple proteins are embedded with the bilayer there are a number of different factors that will influence how they distribute throughout the membrane. The arrangements of proteins in the bilayer could play an important role in a variety of phenomena associated with aggregation and disaggregation of structures. Among the factors that influence how proteins distribute are the direct protein-protein steric interactions, effects mediated for the protein induced deformations of the bilayer by the elastic mechanics of the membrane, and entropic effects driven by the thermal fluctuations of the membrane and proteins. To study the protein distribution within the membrane we consider the radial distribution function given by [5]

$$g(r) = \frac{L^2}{N}\frac{1}{2\pi r}\left\langle \sum_k \delta\left(|\mathbf{r}_k| - r\right)\right\rangle_{\mathbf{r}_0} \tag{23}$$

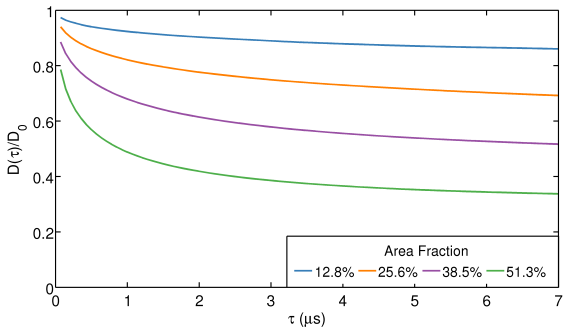

FIG. 5: Diffusivity of a Protein Subject to Steric Interactions. Shown is the effective time-dependent diffusion coefficient $D(\tau)$ of a protein as the total area fraction of proteins in the membrane is varied. Over short times the diffusivity is close to the bare diffusivity of the protein. Over longer time scales the steric interactions inhibit motion of the protein and the effective diffusivity is reduced.

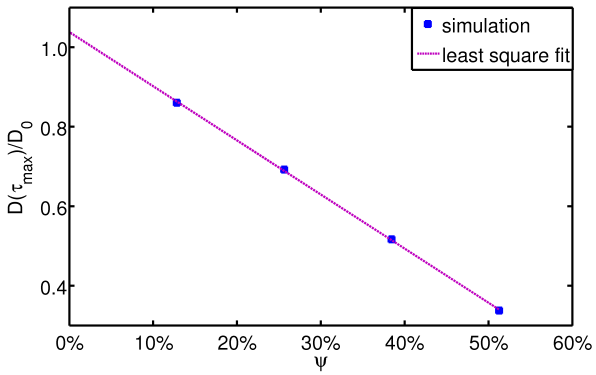

FIG. 6: Diffusivity of a Protein Subject to Steric Interactions. The effective diffusivity is found to be significantly reduced over sufficient time scales by steric interactions with neighboring proteins that restrict protein motions. Shown is the effective diffusion coefficient at time $\tau = \tau_{\max} = 7\mu s$. Interestingly, the diffusivity is found to decrease linearly as the area fraction is varied.

where the $\langle\cdot\rangle_{\mathbf{r}_0}$ denotes an ensemble average conditioned on fixing the location of a given protein at $\mathbf{r}_0$. In practice, we estimate the radial distribution function by considering the number of proteins N_n in successive annuli $\{\mathbf{x}|d_n \leq |\mathbf{x}-\mathbf{r}_0| < d_{n+1}\}$ of area A_n with $d_n = n\delta r$ and using the estimate $g_n = \frac{L^2}{N}\frac{N_n}{A_n}$ and $g(n\delta r) \approx g_n$.

In our initial investigations we consider the radial distribution function for an ensemble of flat proteins within the membrane for various area fractions. These results are shown in Figure 7. We find as the area fraction increases from 12% to 52% that significant structures appear in the relative arrangements of the proteins. This is reflected in the radial distribution function. Some additional insight into the collective interactions that are driving the observed distribution of proteins within the membrane can be obtained by considering the potential of mean force [5] given by

$$A(r) = -k_B T \ln(g(r)). \tag{24}$$

This provides equivalent information as the radial distribution function but highlights the contributions on the energy scale of $k_B T$ and the average forces experienced by a protein at different separation radii from a given protein. The potential of mean force as the area fraction is varied is shown in Figure 8.

To investigate effects arising from the coupling of the proteins with the membrane mechanics, we simulate as a baseline proteins on a flat surface without any membrane mechanics and subject only to the steric interactions. The potential of mean force in this case is shown in Figure 9. To quantitate the contributions made by the membrane mechanics and the membrane fluctuations, we subtract the pure steric potential of mean force from the total potential of mean force to obtain $A_{\mathrm{m}}(r) = A(r) - A_{\mathrm{pp}}(r)$. This remaining membrane associated contribution to the potential of mean force is shown in Figure 10.

We see in in this case that the sterics account for most of the features of the full potential of mean force. Interestingly, we find that as the flat proteins come into close proximity the membrane mechanics and fluctuations act to drive the

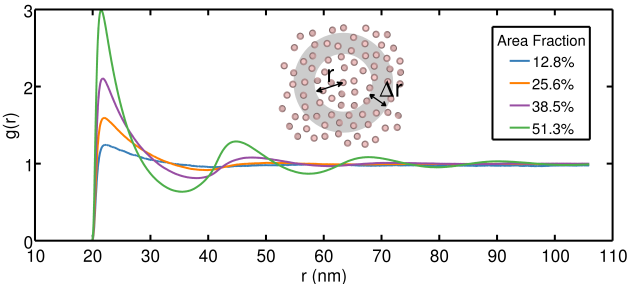

FIG. 7: Radial Distribution Function $g(r)$. Flat proteins diffusing on a fluctuating membrane surface and coupled to the membrane mechanics were simulated with area fractions varied over the range of 12% to 52%. This corresponds to 25 to 100 proteins within the simulation domain. As the area fraction increases the proteins exhibit to an increasing degree an ordered packing structure within the membrane sheet.

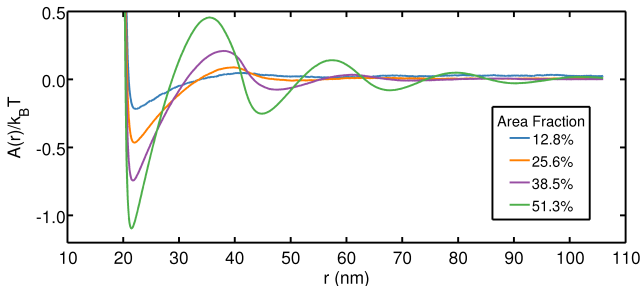

FIG. 8: Potential of Mean Force $A(r)$. Flat proteins diffusing on a fluctuating membrane surface and coupled to the membrane mechanics were simulated with area fractions varied over the range of 12% to 52%. This corresponds to 25 to 100 proteins within the simulation domain. As the area fraction increases the proteins exhibit to an increasing degree an ordered packing structure within the membrane sheet.

proteins together, see for small r Figure 10. This is in contrast to the steric interactions that serve to push the proteins apart when in close proximity. We have found that for systems of flat proteins in which the sterics are neglected, the membrane fluctuations tend to drive proteins to co-locate in space. While a signature of this effect appears, it is rather weak in this case only contributing on the order of 1% of a k_BT to the potential of mean force. These results indicate that the sterics are responsible primarily for the observed radial distribution function and potential of mean force.

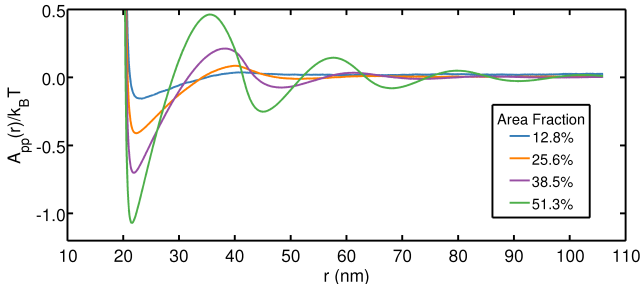

FIG. 9: Potential of Mean Force $A_{\rm pp}(r)$. The potential of mean force for proteins diffusing on a flat static surface without coupling to the membrane mechanics. The proteins are subject only to the steric interactions $\mathcal{H}_{\rm pp}$.

E. Homogenized Elasticity of Membranes with Embedded Proteins

The embedding of proteins within the membrane at sufficient area fractions is expected to augment significantly the membrane elasticity relative to a pure lipid bilayer. An interesting question is how the local augmentation of

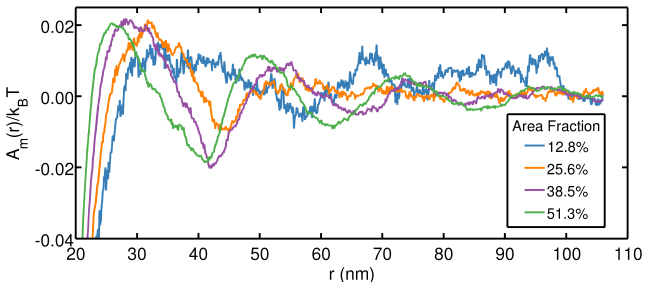

FIG. 10: Potential of Mean Force $A_{\rm m}(r)$. Contributions arising from the protein coupling to the membrane mechanics to the full potential of mean force $A(r) = A_{\rm pp}(r) + A_{\rm m}(r)$.

mechanical properties of the bilayer influence the overall elastic bending modulus of the membrane. We study here how rigid flat proteins embedded within the bilayer augment the bending elasticity as the area fraction is varied.

To obtain the effective homogenized bending elasticity of the heterogeneous membrane-protein sheet, we perform spectral analysis of the fluctuations of the system in the Fourier basis. By comparing the fluctuations of the leading order Fourier modes with that of a purely elastic sheet provides a method to estimate the homogenized bending modulus. Using as our model for a pure elastic sheet the Helfrich free energy and Monge-Gauge, we have for the covariance of the Fourier modes

$$\sigma_q^2 = \langle |\hat{h}_{\mathbf{q}}|^2 \rangle = \frac{1}{K_m} k_B T L^2 \frac{1}{q^4}. \tag{25}$$

An estimate of the effective bending modulus can be obtained from the spectrum of fluctuations of the heterogeneous membrane-protein sheet by fitting the covariances over a range of Fourier modes by adjusting the parameter K_m.

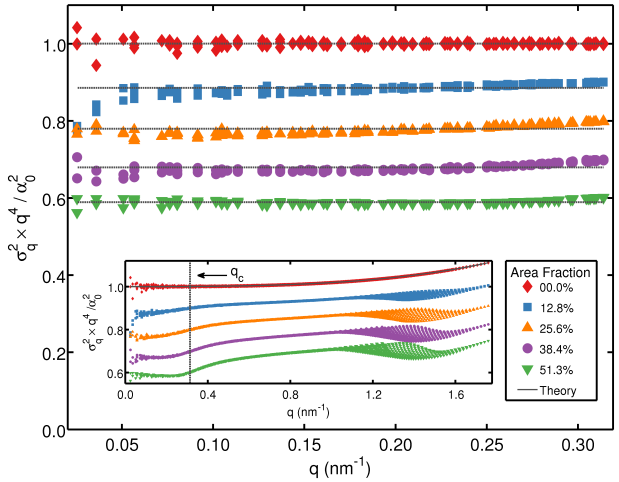

FIG. 11: Scaled Covariances of the Fourier Modes for Membrane Fluctuations. The covariance of each Fourier mode of the fluctuations of the heterogeneous membrane-protein sheet are shown. To more clearly show features of the covariances they have been scaled by q^4. To estimate the effective homogenized bending modulus equation 25 was fit to the covariances for $q < q_c$. Good agreement with the homogenization theory was found over all area fractions considered for Fourier modes having wavelengths smaller than the size of a protein.

An important issue is that Fourier modes are expected to follow the pure elastic sheet trends to a good approximation only over modes for which the wavelengths are significantly larger than the protein size. To characterize where the wavelengths of the Fourier modes become comparable to the protein size, we define the critical wavenumber $q_c = \pi/R_p$ where R_p is given in Table VI. We then perform fits of the Fourier modes only for wave-numbers $q < q_c$. Using this approach we have obtained effective moduli for the homogenized elasticity over a range of protein area fractions. Our results are shown in Figure 11 and Figure 12. The simulations make use of the parameters in Table VI.

The homogenized theory for the spectrum of fluctuations given by equation 25 is found to provide a particularly good fit to the observed fluctuations of the protein-membrane system for $q < q_c$. A particular feature in the analysis

is the scaling of σ_q^2 by q^4 which is found to yield scaled covariances $\sigma_q^2 q^4$ that are well approximated by a constant determined by the single fitting parameter K_m, see Figure 11. This feature of the observed covariances indicates that the membrane-protein system behaves at large length scales to a good approximation like a purely elastic sheet. The covariances for $q > q_c$ depend on features of the membrane-protein sheet at length scales smaller than the protein size and deviate from a purely elastic sheet as should be expected.

An interesting issue for membrane-protein sheet is how the bending elasticity depends on the area fraction of proteins. A mixture theory for how a heterogeneous elastic sheet that mixes two different elastic bending moduli at different area fractions was developed in [24]. For our membrane-protein model this corresponds to the bending elasticity purely of the bilayer of K_m and the bending elasticity in the area occupied by the protein of K_p. The mixture theory predicts that for such a heterogeneous elastic sheet the effective bending modulus K_* depends on the area fraction as

$$\frac{1}{K_*} = \frac{1-\psi}{K_m} + \frac{\psi}{K_p}. \tag{26}$$

The area fraction is given by $\psi = NA'_p/L^2$, where N is the number of proteins and A'_p is the effective area per protein within the bilayer sheet. We compare this theory with the effective homogenized bending modulus obtained from equation 25 in Figure 12. We find very good agreement with theory when the effective protein area is $A'_p = R_p^2\pi$. The theory performs very well for area fractions less than 30%. At larger area fractions we find the effective bending modulus of the homogenized membrane-protein sheet deviates slightly, see Figure 12. At such area fractions the simple mixture model is expected to break down. The particular interactions and arrangements of the proteins in coordination with the bending modes of the sheet are expected to play a more significant role in contributing to the homogenized elasticity.

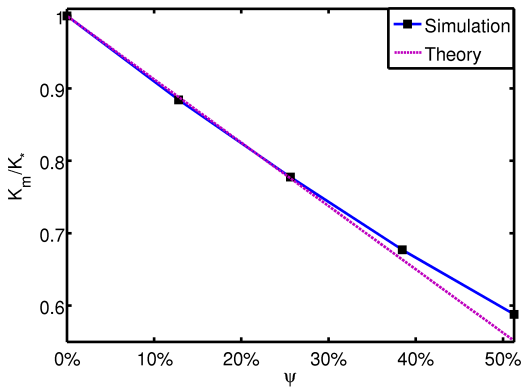

FIG. 12: Effective Bending Modulus of Homogenized Protein-Membrane Sheet. The homogenized bending elasticity of a membrane-protein sheet are shown as the area fraction of proteins is varied. The membrane contains flat proteins that locally make the bilayer rigid with bending modulus K_p. The mixture theory for effective bending elasticity given in equation 26 is compared with homogenized elasticity of the membrane-protein sheet. Good agreement is found for area fractions less than 30%. For larger area fraction slight deviations occur likely arising from important collective interactions among the proteins that are neglected in the mixture model.

F. Membrane-Protein Sheet Fluctuations for $q > q_c$

The physical analysis presented has focused on phenomena for which the most relevant Fourier modes where $q < q_c$. However, for $q > q_c$ some interesting features manifest in the covariance spectrum of the simulated membrane-protein system, see the inset of Figure 11. We discuss briefly the source of these features of the covariance spectrum and their possible relevance to phenomena associated with particular arrangements of the proteins, choice of underlying kernel for the protein model, and the computational methods. For a given fixed configuration of the proteins, the full covariance of the Fourier modes of the membrane-protein system are given by

$$\langle \hat{\mathbf{h}}\hat{\mathbf{h}}^T \rangle = k_B T L^2 \widehat{B}^{-1}. \tag{27}$$

The $\mathbf{h}$ gives the composite vector of Fourier modes with $h_{\mathbf{q}} = [\mathbf{h}]_{\mathbf{q}}$. The $\widehat{B}$ represents in Fourier space the linear force response of the membrane to deformations, $\mathbf{F}[h] = -Bh$ with $\widehat{B} = \mathcal{F}B\mathcal{F}^{-1}$ where $\mathcal{F}_{\mathbf{q},\mathbf{m}} = (L/M)^2 e^{-i(\mathbf{m}\cdot\mathbf{q})L/M}$

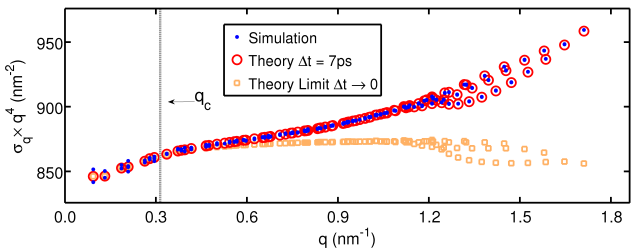

FIG. 13: Theory for High Frequency Fourier Modes. The covariance structure of an inhomogeneous membrane-protein sheet manifests significant deviations from the covariance structure of a pure Helfrich elastic sheet. Theory based on the linear force response of the membrane for a fixed arrangement of the proteins is found to predict well the observed covariance structure from simulations.

represents the Fourier transform change of basis. For a homogeneous membrane subject purely to the Helfrich elastic bending energy, the force response operator is diagonal in Fourier space with

$$\widehat{B}_{\mathbf{p},\mathbf{q}} = \begin{cases} K_m q^4 & \mathbf{q} = \mathbf{p}, \\ 0 & \mathbf{q} \neq \mathbf{p}, \end{cases} \tag{28}$$

In practice, the simulations exhibit deviation from this covariance structure in the high frequency regime $q > q_c$ even for the homogeneous membrane. This highlights an important consideration when performing such simulations requiring in the computational methods control of such deviations. These deviations can be explained readily by a theory based on considering the discrete time stochastic dynamics inherent when using the Euler-Maruyama numerical method with time step Δt. For the discrete time stochastic system the covariance structure is predicted to have

$$\langle \hat{\mathbf{h}}\hat{\mathbf{h}}^T \rangle = k_B T L^2 \left(\widehat{B} - \frac{\Delta t}{2} \Lambda \widehat{B}^2 \right)^{-1}. \tag{29}$$

This follows from the discrete time dynamics given by equation 13 which yields the following equation for the moments

$$\langle \hat{\mathbf{h}}^{n+1}\hat{\mathbf{h}}^{n+1,T} \rangle = \left(\mathcal{I} - \Delta t \Lambda \widehat{B} \right) \langle \hat{\mathbf{h}}^{n}\hat{\mathbf{h}}^{n,T} \rangle \left(\mathcal{I} - \Delta t \Lambda \widehat{B} \right)^T + \Delta t k_B T L^2 \Lambda \langle \hat{\xi}^n \hat{\xi}^{n,T} \rangle. \tag{30}$$

By using the properties of $\widehat{B}$ and $\langle \hat{\mathbf{h}}\hat{\mathbf{h}}^T \rangle \left(\mathcal{I} - \Delta t \Lambda \widehat{B} \right)^T = \left(\mathcal{I} - \Delta t \Lambda \widehat{B} \right) \langle \hat{\mathbf{h}}\hat{\mathbf{h}}^T \rangle$, we obtain equation 29.

We show that this theory precisely predicts the observed covariance structure of a homogeneous membrane, see theory line on inset of Figure 11. In particular, the theory predicts rather precisely the deviations that occur for the high frequency Fourier modes $q > q_c$. This theory has important implications for numerical simulations requiring that a sufficiently small time step be chosen not only to ensure stability of the numerical methods but also to ensure accuracy so that the physics of the covariance structure is captured correctly. The characteristic timescale for a given mode $\mathbf{q}$ of the model is given by

$$\tau_{\mathbf{q}} = \frac{2\eta}{K_m} q^{-3}. \tag{31}$$

This can be used to characterize in the model the fastest timescale $\tau_f = \min\{\tau_{\mathbf{q}}\}$ and the slowest timescale $\tau_s = \max\{\tau_{\mathbf{q}}\}$ associated with the dynamics. The deterministic homogeneous version of equation 13 integrated numerically with Euler-Marayuma has the stability constraint

$$\frac{\Delta t}{\tau_s} \leq \frac{\tau_f}{\tau_s} = \left(\frac{q_f}{q_s} \right)^{-3}. \tag{32}$$

We used this upper bound for the purely homogeneous membrane as a guide when choosing our timestep for the membrane-protein system. We found that in practice we needed a time step of at most $\Delta t < 10$ps to ensure there were not any significant deviations in the Fourier modes above the critical wavenumber $q > q_c$.

For inhomogeneous membranes the covariance structure manifests significant scatter for the high frequency modes, indicating breaking of rotational symmetry of the system. We show this feature arises from the discrete nature of the proteins and their particular arrangements within the bilayer. To illustrate the contributions of these effects we again consider the linear force response of the membrane $\widehat{B}$ for a fixed arrangement of the proteins. The operator $\widehat{B}$ now

includes in addition to the Helfrich bending elasticity contributions from the proteins on the membrane mechanics, see equation 2. We find that the theory developed in equation 29 predicts well the covariance structure of fluctuations now of the heterogeneous membrane-protein sheet, see Figure 13. These results indicate that the scatter observed is a consequence of the breaking of rotational symmetry by the discrete nature of the proteins and their particular arrangement. Depending on the physical phenomena under study this theory could be useful to help inform an appropriate choice of kernel function and in the development of appropriate numerical methods to reliably model and simulate a given membrane-protein system.

V. CONCLUSION

We introduced new methods for the investigation of proteins embedded within lipid bilayer membranes based on an approach closely related to the *Immersed Boundary Method* [1, 27]. Our approach captures phenomena at the level of individual proteins and their contributions to membrane mechanics. We also capture for the membrane the hydrodynamic coupling and the thermal fluctuations. We have presented analysis concerning the statistical mechanics of our model. We have established some rather general guidelines for introducing approximations in a manner consistent with statistical mechanics that should be useful in the development of related models. We have also developed computational methods to overcome related issues with respect to finite lattice effects associated with the discretization of the membrane-protein equations, which we call *Phase Factor Averaging*. We showed how these methods can be used to further investigate the effective diffusivity of curvature inducing proteins embedded within a fluctuating bilayer. We also showed that the methods can be used to study the influence of the protein mechanics on the effective bending elasticity of bilayer sheets as the area fraction is varied. In summary, the presented modeling approach and computational methods are quite general and should be useful in the study of a wide variety of phenomena relevant to proteins embedded within lipid bilayer membranes.

VI. ACKNOWLEDGEMENTS

The author P.J.A acknowledges support from research grant NSF CAREER - 0956210. We also acknowledge support from the UCSB Center for Scientific Computing NSF MRSEC (DMR-1121053) and UCSB MRL NSF CNS-0960316. For stimulating discussions and helpful suggestions, we would also like to thank Philip Pincus, Markus Deserno, Brian Camley, and Sveinn Thorarinsson. We also participated and benefited from two helpful workshops during the writing of this paper held at the Kavli Institute for Theoretical Physics China, CAS, Beijing 100190, China and the Kavli Institute for Theoretical Physics, Santa Barbara, California, USA. National Science Foundation under Grant No. NSF PHY05-51164.

[1] P. J. Atzberger, P. R. Kramer, and C. S. Peskin, *A stochastic immersed boundary method for fluid-structure dynamics at microscopic length scales*, Journal of Computational Physics, 224 (2007), pp. 1255–1292–.
[2] J.Lewis M.Raff K.Roberts B.Alberts, A.Johnson and P.Walker, *Molecular Biology of the Cell*, Garland Publishing, 2002.
[3] Frank L.H. Brown, *Elastic modeling of biomembranes and lipid bilayers*, Annu. Rev. Phys. Chem., 59 (2008), pp. 685–712.
[4] Richard L. Burden and Douglas Faires, *Numerical Analysis*, Brooks/Cole Cengage Learning, 2010.
[5] D. Chandler, *Introduction to Modern Statistical Mechanics*, Oxford, New York, 1987.
[6] Kung Ching, *A possible unifying principle for mechanosensation*, Nature, 436 (2005), pp. 647–54–.
[7] James W. Cooley and John W. Tukey, *An algorithm for the machine calculation of complex fourier series*, Math. Comput., 19 (1965), p. 297–301.
[8] Gary J. Doherty and Harvey T. McMahon, *Mediation, modulation, and consequences of membrane-cytoskeleton interactions*, Annu. Rev. Biophys., 37 (2008), pp. 65–95.
[9] M. D. Edwards, I. R. Booth, and S. Miller, *Gating the bacterial mechanosensitive channels: Mscs a new paradigm?*, Current Opinion in Microbiology, 7 (2004), pp. 163–167–.
[10] Donald M. Engelman, *Membranes are more mosaic than fluid*, Nature, 438 (2005), pp. 578–580.
[11] Daan Frenkel and Berend Smit, *Understanding Molecular Simulation*, Academic Press, San Diego, second ed., 2002. pp. 525-532.
[12] C. W. Gardiner, *Handbook of stochastic methods*, Series in Synergetics, Springer, 1985.
[13] I. M. Gelfand and S. V. Fomin, *Calculus of Variations*, Dover, 2000.

[14] J. Gullingsrud and K. Schulten, *Lipid bilayer pressure profiles and mechanosensitive channel gating*, Biophys. J., 86 (2004), pp. 3496–3509–.
[15] W. Helfrich, *Elastic properties of lipid bilayers: theory and possible experiments*, Z. Nat., 28 (1973), pp. C693–703.
[16] J. Howard, *Mechanics of motor proteins and the cytoskeleton*, Sinauer, Sunderland, Massachusetts, 2001.
[17] Richard Hynes, *Integrins: Bidirectional, allosteric signaling machines*, Cell, 110 (2002), pp. 673–687.
[18] Ken Jacobson, Erin D. Sheets, and Rudolf Simson, *Revisiting the fluid mosaic model of membranes*, Science, 268 (1995), pp. 1441–1442.
[19] Tomas Kirchhausen, *Clathrin*, Annu. Rev. Biochem., 69 (2000), pp. 699–727.
[20] A. Kusumi and Y. Sako, *Cell surface organization by the membrane skeleton*, Curr. Opin. Cell Biol., 8 (1996), pp. 566–574.
[21] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics*, Butterworth Heineman, Oxford, UK, third ed., 1986.
[22] Alex J. Levine and F. C. MacKintosh, *Dynamics of viscoelastic membranes*, Phys. Rev. E, 66 (2002), pp. 061606–.
[23] Ali Naji, Paul J. Atzberger, and Frank L. H. Brown, *Hybrid elastic and discrete-particle approach to biomembrane dynamics with application to the mobility of curved integral membrane proteins*, Phys. Rev. Lett., 102 (2009), pp. 138102–.
[24] Roland R. Netz and P. Pincus, *Inhomogeneous fluid membranes: Segregation, ordering, and effective rigidity*, Phys. Rev. E, 52 (1995), pp. 4114–.
[25] B. Oksendal, *Stochastic Differential Equations: An Introduction*, Springer, 2000.
[26] J.C. Owicki and H. M. McConnell, *Theory of protein-lipid and protein-protein interactions in bilayer membranes*, Proc. Natl. Acad. Sci, 76 (1979), pp. 4750–4754.
[27] C. S. Peskin, *The immersed boundary method*, Acta Numerica, 11 (2002), pp. 479–517.
[28] Rob Phillips, Tristan Ursell, Paul Wiggins, and Pierre Sens, *Emerging roles for lipids in shaping membrane-protein function*, Nature, 459 (2009), pp. 379–385.
[29] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, Cambridge University Press, Cambridge, 1994.
[30] L. E. Reichl, *A Modern Course in Statistical Physics*, John Wiley and Sons, 1998.
[31] Ellen Reister-Gottfried, Stefan M. Leitenberger, and Udo Seifert, *Hybrid simulations of lateral diffusion in fluctuating membranes*, Phys. Rev. E, 75 (2007), p. 011908.
[32] H. Royden, *Real Analysis*, Simon & Schuster Company, 1988.
[33] S. J. Singer and Garth L. Nicolson, *The fluid mosaic model of the structure of cell membranes*, Science, 175 (1972), pp. 720–731.
[34] S. Sukharev and A. Anishkin, *Mechanosensitive channels: what can we learn from 'simple' model systems?*, Trends in Neurosciences, 27 (2004), pp. 345–351–.
[35] M. E. Tuckerman, C. J. Mundy, and G. J. Martyna, *On the classical statistical mechanics of non-hamiltonian systems*, EPL (Europhysics Letters), 45 (1999), pp. 149–155.

Appendix A: Statistical Mechanics of Membrane-Protein System

An important feature of the membrane-protein system is the coupled motions and simultaneous fluctuations of the protein and bilayer membrane. This can result in subtle issues with respect to formulating appropriate stochastic dynamics consistent with statistical mechanics. To cope with this issue, we perform systematic analysis to obtain conditions useful to guiding such modeling. We derive particular conditions that ensure the stochastic dynamics satisfy the fluctuation-dissipation principle and have as their equilibrium distribution the Gibbs-Boltzmann distribution with detailed balance.

For the purpose of this analysis, it is convenient to rewrite the membrane-protein stochastic dynamics in terms of a grand mobility tensor as

$$\frac{\partial \mathbf{Z}}{\partial t} = M(-\nabla_Z \mathcal{H}) + \mathbf{w} + \mathbf{g}. \tag{A1}$$

In this notation $\mathbf{Z} = [h, \mathbf{r}]$ denotes the composite vector of all degrees of freedom of the membrane-protein system, namely, the membrane configuration and all protein positions. The $\nabla_Z \mathcal{H} = [\delta\mathcal{H}/\delta h, \nabla_{\mathbf{r}}\mathcal{H}]$ denotes the composite vector of all derivatives of the Hamiltonian. The $\mathbf{g}$ denotes the stochastic terms accounting for thermal fluctuations. We shall assume throughout that $\mathbf{g}$ is Gaussian with mean 0 and covariance $G = \langle \mathbf{g}\mathbf{g}^T \rangle$. The $\mathbf{w}$ drift term arises from possible variability of the strength of the thermal fluctuations that depend on the state of the system and the reduction of the inertial dynamics to the overdamped regime [12, 35]. The grand mobility tensor can be expressed as

$$M = \begin{bmatrix} M_{hh} & M_{hr} \\ M_{rh} & M_{rr} \end{bmatrix}. \tag{A2}$$

The components $M_{\alpha\beta}$ denote the mobilities of the membrane and proteins giving the rate of change of the configuration in response to forces. In particular, $M_{\alpha\beta}$ gives the overdamped velocity response of α to a force acting on β, where $\alpha, \beta \in \{h, r\}$ indicates either the membrane or proteins.

Depending on the discussion, we shall also find it convenient to express the dynamics as

$$\frac{dh}{dt} = M_{hh}\left(-\frac{\delta \mathcal{H}}{\delta h}\right) + M_{hr}\left(-\nabla_{\mathbf{r}}\mathcal{H}\right) + \mathbf{w}_h + \mathbf{F}_{\text{thm}} \tag{A3}$$

$$\frac{d\mathbf{r}}{dt} = M_{rr}\left(-\nabla_{\mathbf{r}}\mathcal{H}\right) + M_{rh}\left(-\frac{\delta \mathcal{H}}{\delta h}\right) + \mathbf{w}_r + \mathbf{f}_{\text{thm}} \tag{A4}$$

where the $\mathbf{F}_{\text{thm}}, \mathbf{f}_{\text{thm}}$ correspond to the terms accounting for the thermal fluctuations.

An important consideration for a proposed stochastic dynamics is whether or not it has as its equilibrium distribution the Gibbs-Boltzmann distribution with detailed balance. The Gibbs-Boltzmann distribution must remain invariant under the dynamics in a manner consistent with microscopic time-reversibility of the physical process. In the overdamped regime, the principle of detailed balance requires the Gibbs-Boltzmann distribution to remain invariant under the stochastic dynamics with no-net flux between states. Mathematically, this condition is expressed by considering the Fokker-Planck equation for equation A1 given by [12, 25]

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot J \tag{A5}$$

$$J = \left[M\left(-\nabla \mathcal{H}\right) + \mathbf{w} - \frac{1}{2}\nabla \cdot G\right]\rho - \frac{1}{2}G\nabla\rho. \tag{A6}$$

The drift term $-\frac{1}{2}\nabla \cdot G$ arises as a consequence of the stochastic term possibly having a strength that depends on the current state of the system $G = G(\mathbf{z})$ with $\mathbf{z} = [h, \mathbf{r}]$, see [12, 25].

At thermodynamic equilibrium, the Gibbs-Boltzmann distribution must be invariant with detailed balance under the stochastic dynamics so that $J = 0$. The Gibbs-Boltzmann distribution for the membrane-protein system is given by

$$\rho(\mathbf{z}) = \frac{1}{\mathcal{Z}}\sqrt{\mathfrak{g}}e^{-\mathcal{H}/k_B T}. \tag{A7}$$

The $\mathcal{H} = \mathcal{H}(\mathbf{z}) = \mathcal{H}[h, \mathbf{r}]$ is the Hamiltonian, k_B is Boltzmann's constant, T is temperature, $\mathcal{Z}$ is the partition function that formally normalizes the density to one [30]. The $\mathfrak{g} = \mathfrak{g}[\mathbf{z}] = \mathfrak{g}[h, \mathbf{r}]$ is the induced metric factor on phase-space arising from the use of generalized coordinates [35].

For the membrane-protein system the factor $\mathfrak{g}$ arises from the use of the Monge-Gauge and the protein location projected from the membrane surface to the xy-plane. In particular, this projection contributes from the membrane geometry the factor $\mathfrak{g} = \prod_k g^{(k)}$ with $g^{(k)} = 1 + \left(\nabla h\left(\mathbf{r}_k\right)\right)^2$, where $\mathbf{r}_k$ is the location of the k^{th} protein. We also remark that technically the density function ρ is only formal given the mathematical issues associated with the representation of measures on function spaces (Radon - Nikodym derivative [32]). The density ρ is useful none-the-less to convey the main results which are obtainable from a truncated series approximation or more rigorous approaches [25, 32].

Detailed balance $J = 0$ corresponds to the following condition after substitution of equation A7 into equation A6

$$\left(\mathbf{w} - \frac{1}{2}\nabla \cdot G - \frac{1}{2}G\nabla \log(\sqrt{\mathfrak{g}})\right) - \left(M - \frac{1}{2k_B T}G\right)\nabla \mathcal{H} = 0. \tag{A8}$$

Since this must hold for all possible choices for the Hamiltonian $\mathcal{H}$, the condition of detailed balance requires

$$G = 2k_B T M \tag{A9}$$

$$\mathbf{w} = \frac{1}{2}\nabla \cdot G + \frac{1}{2}G\nabla \log(\sqrt{\mathfrak{g}}). \tag{A10}$$

These conditions arise using the arbitrariness of $\mathcal{H}$ by first setting $\mathcal{H} = 0$ to obtain equation 7 which upon substitution and then setting $\mathcal{H}$ so $\nabla \mathcal{H} \neq 0$ then yields equation 6. The first condition given by equation 6 corresponds to the fluctuation-dissipation principle of statistical mechanics [30]. The second condition arises from the generalized coordinates and state-dependent intensity of the contributions of the thermal fluctuations to the dynamics.

In the special case when $G = G_0$ and $\mathfrak{g} = \mathfrak{g}_0$ are both constant, the second condition simply yields $\mathbf{w} = 0$ and the fluctuation-dissipation principle is found sufficient to ensure detailed balance. In the case that either $G = G(\mathbf{z})$ or $\mathfrak{g} = \mathfrak{g}(\mathbf{z})$ are non-constant, the thermally induced drift term $\mathbf{w}$ may become non-zero. In the case $\mathfrak{g}$ is non-constant, prior models have often introduced a metric term in the mobility tensor for the force response of the proteins. If that term indeed contributes significantly, one expects that the often neglected thermal drift term would also contribute significantly to the dynamics. Overall, these drift terms arise from the variability of the strength of the thermal

fluctuations that depend on the state of the system and the reduction of the inertial dynamics to the overdamped regime [12, 35].

The conditions of equation 6 and equation 7 provide useful guidance when formulating models of the membrane-protein stochastic dynamics to ensure consistency with statistical mechanics. Unfortunately, many prior models proposed for the membrane-protein stochastic dynamics do not satisfy these conditions, as we discuss in Section II B 1.

Appendix B: Numerical Approximation of the Forces

The discretization of the force density acting on the membrane $\mathbf{F_q} = -\frac{\delta \mathcal{H}}{\delta h}$ is obtained spectrally by applying the phase factor averaging Fourier transform in equation 18 to the variational derivative of the Hamiltonian i.e.

$$\begin{aligned}\mathbf{F_q} = [\bar{\mathcal{F}}\mathbf{F}]_\mathbf{q} = -K_m q^4 h_\mathbf{q} + q^2 \left[\bar{\mathcal{F}}\Big(G_p(\mathbf{x}-\mathbf{r}_k)[(K_p-K_m)\nabla^2 h - 2K_p C_p]\Big)\right]_\mathbf{q} \\ +q_1^2\left[\bar{\mathcal{F}}\Big(G_p(\mathbf{x}-\mathbf{r}_k)(K_p'-K_m')\partial_{yy}h\Big)\right]_\mathbf{q} + q_2^2\left[\bar{\mathcal{F}}\Big(G_p(\mathbf{x}-\mathbf{r}_k)(K_p'-K_m')\partial_{xx}h\Big)\right]_\mathbf{q} \\ -2q_1q_2\left[\bar{\mathcal{F}}\Big(G_p(\mathbf{x}-\mathbf{r}_k)(K_p'-K_m')\partial_{xy}h\Big)\right]_\mathbf{q}. \end{aligned} \tag{B1}$$

We must approximate not only on the values of the height function h at the lattice sites but also the second derivatives of h. For this purpose, we use the Fourier transform to obtain a spectral approximation for the derivatives of h. When also using phase factor averaging ($N_a > 1$) this also requires obtaining values for h and the second derivatives of h at the new shifted lattice locations which we obtain by spectral interpolation. In our initial approach we use the Trapeziodal rule to approximate the force acting on the protein which for periodic functions gives the discrete approximation

$$\mathbf{f}_k = \nabla_{\mathbf{r}_k}\mathcal{H} \approx \frac{1}{2}\sum_{\mathbf{m}}\left(\frac{1}{N_s}\sum_{l=1}^{N_s}\nabla_{\mathbf{r}_k} Q[h,\mathbf{r}](\mathbf{x_m}-\mathbf{s}_l)\right)\Delta\mathbf{x}, \tag{B2}$$

where $\mathbf{x_m}$ denotes the lattice sites with index $\mathbf{m} = (m_1, m_2)$ and $\Delta\mathbf{x}$ denotes the lattice spacing, $\mathbf{s}_l$ denotes the shift vectors and $Q[h,\mathbf{r}]$ is the energy density associated with height profile h and protein positions $\mathbf{r}$. The required values of the height function h have already been computed when obtaining the force density in equation B1. We remark that while this yields good results for the derivatives, the dominant source of error in the discretization will be at the level of the quadrature used to approximate the energy and subsequently derived forces that act on the proteins. This could be improved by the use of more sophisticated quadrature methods but at additional computational expense.

Appendix C: Kernel Function Definitions

For the kernel function $G_p(\mathbf{x})$ that couples the protein and the membrane we use the following form

$$G_p(\mathbf{x}) = \frac{A_p}{a^2}\phi(x_1/a)\phi(x_2/a), \tag{C1}$$

where A_p (we take $A_p = 100\text{nm}^2$) denotes an effective area to attribute to the protein, and $8a$ the protein size (we take $a = 2.5$ nm). In our investigations of the numerical methods we considered three different choices for the function ϕ, which we refer to as (i) Hat function ϕ_h, (ii) Gaussian function ϕ_g, (iii) Peskin function ϕ_p. The Hat function is given by

$$\phi_h(u) = (1 - |u/4|)/4 \tag{C2}$$

where $|u| < 4$ and $\phi_h(u) = 0$ for $|u| \geq 4$. The Gausssian function is given by

$$\phi_g(u) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma}e^{\frac{-u^2}{2\sigma^2}} & -4 \leq u < 4, \\ 0 & \text{else}, \end{cases} \tag{C3}$$

where $\sigma = 4/3$. The Peskin function is given by

$$\phi_p(u) = \frac{1}{16}\begin{cases} 3 - u + \sqrt{1+2u-u^2} & 0 \leq u \leq 2, \\ 5 - u - \sqrt{-7+6u-u^2} & 2 \leq u \leq 4, \\ 0 & 4 \leq u. \end{cases} \tag{C4}$$

For $u < 0$, we use $\phi_p(-u) = \phi_p(u)$ which yields an even extension of ϕ_p for all u.