

## Coupling all-atom molecular dynamics simulations of ions in water with Brownian dynamics

Dr. M Kothandaraman, G Madhusudhana, B Murlikrishna

<sup>1</sup>Professor, <sup>2,3</sup>Assistant Professor

Department of Maths

Bheema Institute of Technology and Science, Adoni

### ABSTRACT:

Investigations are conducted using molecular dynamics (MD) simulations of ions (K<sup>+</sup>, Na<sup>+</sup>, Ca<sup>2+</sup>, and Cl<sup>-</sup>) in aqueous solutions. We use the SPC/E model to explain water. Using MD simulations, a stochastic coarse-grained description of ion behavior is presented and parameterized. It describes the position, acceleration, and velocity of the ions and is presented as a system of coupled stochastic and ordinary differential equations. A description in between Brownian dynamics (BD) models and all-atom MD simulations is offered by the stochastic coarse-grained model. It is applied to the development of a multiscale technique that employs (less detailed) BD simulations in the remaining computational region and all-atom MD simulations in other areas.

**Key words:** Brownian dynamics, molecular dynamics, and multiscale modeling

### 1 Introduction:

Molecular dynamics (MD) simulations of ions in aqueous solutions are limited to modelling processes in relatively small domains containing (only) several thousands of water molecules [1, 2]. Ions play important physiological functions in living cells which typically consist of 10<sup>10</sup>–10<sup>12</sup> water molecules. In particular, processes which include transport of ions between different parts of a cell cannot be simulated using standard all-atom MD approaches. Coarser models are instead used in applications. Examples include Brownian dynamics (BD) simulations [3] and mean-field Poisson-Nernst-Planck equations [4]. In BD methods, individual trajectories of ions are described using

$$dX_i = \sqrt{2D} dW_i, \quad i = 1, 2, 3, \quad (1.1)$$

where  $X = [X_1, X_2, X_3]$  is the position of the ion,  $D$  is its diffusion constant and  $W_i$ ,  $i = 1, 2, 3$ , are three independent Wiener processes [5]. BD description (1.1) does not explicitly include solvent molecules in the simulation. Moreover, in applications, equation (1.1) can be discretized using a (nanosecond) time step which is much larger than the typical time step of MD simulations (femtosecond) [6]. This makes BD less computationally intensive than the corresponding MD simulations.

Longer time steps of BD simulations enable efficient simulations of ion transport between different parts of the cell, but they limit the level of detail which can be incorporated into the model. For example, intracellular calcium is regulated by the release of Ca<sup>2+</sup> ions from the endoplasmic reticulum via inositol-4,5-triphosphate receptor (IP3R) channels. BD models in the literature use equation (1.1) to describe trajectories of calcium ions [3, 7]. The conformational changes between the open and closed states of IP3R channels are controlled

by the binding of Ca<sup>2+</sup> to activating and inhibitory binding sites. BD models postulate that binding of an ion occurs with some probability whenever the distance between the ion and an empty site is less than the specific distance, the so called reaction radius [8, 9]. Although details of the binding process are known [10, 11], they cannot be incorporated into coarse BD models of calcium dynamics, because equation (1.1) does not correctly describe short time behaviour of ion dynamics.

The calcium induced calcium release through IP3R channels is an example of a multiscale dynamical problem where MD simulations are important only in certain parts of the computational domain (close to an IP3R channel), whilst in the remainder of the domain a coarser, less detailed, BD method could be used (to describe trajectories of ions). Such multiscale problems cannot be simulated using MD methods, but there is potential to design multiscale computational methods which compute the desired information with an MD-level of resolution by using MD and BD models in different parts of the computational domain [12].

In [12], three relatively simple and analytically tractable MD models are studied (describing heat bath molecules as point particles) with the aim of developing and analyzing multiscale methods which use MD simulations in parts of the computational domain and less detailed BD simulations in the remainder of the domain. In this follow up paper, the same question is investigated in allatom MD simulations which use the SPC/E model of water molecules. In order to couple MD and BD simulations, we need to first show that the MD model is in a suitable limit described by a stochastic model which does not explicitly take into account heat bath (water) molecules. In [12], this coarser description was given in terms of Langevin dynamics. Considering all-atom MD simulations, the coarser stochastic model of an ion is more complicated than Langevin dynamics. In this paper, it will be given by

$$dX_i = V_i dt, \tag{1.2}$$

$$dV_i = U_i dt, \tag{1.3}$$

$$dU_i = (-\eta_1 V_i + Z_i) dt, \tag{1.4}$$

$$dZ_i = -(\eta_2 Z_i + \eta_3 U_i) dt + \eta_4 dW_i, \quad i=1,2,3, \tag{1.5}$$

where  $X \equiv [X_1, X_2, X_3]$  is the position of the ion,  $V \equiv [V_1, V_2, V_3]$  is its velocity,  $U \equiv [U_1, U_2, U_3]$  is its acceleration,  $Z \equiv [Z_1, Z_2, Z_3]$  is an auxiliary variable,  $dW \equiv [dW_1, dW_2, dW_3]$  is white noise and  $\eta_j, j = 1, 2, 3, 4$ , are parameters. These parameters will be chosen according to all-atom MD simulations as discussed in Section 3. In Section 4, we show that (1.2)–(1.5) provides a good approximation of ion behaviour. In Section 5, we further analyse the system (1.2)–(1.5) and show how parameters  $\eta_j, j = 1, 2, 3, 4$ , can be connected with diffusion constant  $D$  used in the BD model (1.1).

ion	$A_{io}$ [Da Å <sup>14</sup> ps <sup>-2</sup> ]	$B_{io}$ [Da Å <sup>8</sup> ps <sup>-2</sup> ]	$q_i$ [e]	$M$ [Da]
K <sup>+</sup>	$2.8973 \times 10^8$	$2.4587 \times 10^5$	+1	39.0983
Na <sup>+</sup>	$6.6813 \times 10^7$	$1.1807 \times 10^5$	+1	22.9898
Ca <sup>2+</sup>	$1.1961 \times 10^8$	$1.5797 \times 10^5$	+2	40.078
Cl <sup>-</sup>	$1.8038 \times 10^9$	$6.1347 \times 10^5$	-1	35.453

Table 1. Parameters of all-atom MD simulations of ions.

The coarse-grained model (1.2)–(1.5) is used as an intermediate model between the all-atom MD model and BD description (1.1). In Section 5, we show how it can be coupled with the BD model which uses a much larger time step than the MD model. In Section 6, the coarse-grained model (1.2)–(1.5) is coupled with all-atom MD simulations. We then show that all-atom MD models of ions can be coupled with BD description (1.1) using the intermediate coarse-grained model (1.2)–(1.5). We conclude by discussing related methods developed in the literature in Section 7.

### Molecular dynamics simulations of ions in SPC/E water

There have been several MD models of liquid water developed in the literature. The simplest models (for example, SPC [13], SPC/E [14] and TIP3P [15]) include three sites in total, two hydrogen atoms and an oxygen atom. More complicated water models include four, five or six sites [16, 17]. In this paper, we use the threesite SPC/E model of water which was previously used for MD simulations of ions in aqueous solutions [18, 1]. In the SPC/E model, the charges ( $q_h = 0.4238 e$ ) on hydrogen sites are at 1 Å from the Lennard-Jones center at the oxygen site which has negative charge  $q_o = -0.8476 e$ . The HOH angle is 109.47°. We use the RATTLE algorithm [19] to satisfy constraints between atoms of the same water molecule.

We investigate four ions (K<sup>+</sup>, Na<sup>+</sup>, Ca<sup>2+</sup> and Cl<sup>-</sup>) at 25 °C using MD parameters presented in [18]. Let us consider a water molecule and let us denote by  $r_{i0}$  (resp.,  $r_{i1}$  and  $r_{i2}$ ) the distance between the ion and the oxygen site (resp., the first and second hydrogen sites). The pair potential between the water molecule and the ion is then given by [1, 18],

$$A_{io} \left( \frac{1}{r_{io}} \right)^{12} - B_{io} \left( \frac{1}{r_{io}} \right)^6 + k_e \frac{q_i q_o}{r_{io}} + k_e \frac{q_i q_h}{r_{i1}} + k_e \frac{q_i q_h}{r_{i2}}, \quad (2.1)$$

where  $A_{io}$  and  $B_{io}$  are Lennard-Jones parameters between the oxygen on the water molecule and the ion,  $k_e$  is Coulomb's constant and  $q_i$  is the charge on the ion. The values of parameters are given for four ions considered in Table 1. We express mass in daltons (Da), length in ångströms (Å) and time in picoseconds (ps), consistently in the whole paper. Using these units, the parameters of the

ion	$D$ [Å <sup>2</sup> ps <sup>-1</sup> ]	$\langle V_i^2 \rangle$ [Å <sup>2</sup> ps <sup>-2</sup> ]	$\langle U_i^2 \rangle$ [Å <sup>2</sup> ps <sup>-4</sup> ]	$\langle Z_i^2 \rangle$ [Å <sup>2</sup> ps <sup>-6</sup> ]
K <sup>+</sup>	0.183	6.32	$4.86 \times 10^3$	$1.65 \times 10^7$
Na <sup>+</sup>	0.128	10.8	$2.21 \times 10^4$	$8.88 \times 10^7$
Ca <sup>2+</sup>	0.053	6.18	$1.87 \times 10^4$	$9.23 \times 10^7$
Cl <sup>-</sup>	0.177	6.98	$6.56 \times 10^3$	$2.97 \times 10^7$

Table 2. Average values obtained by all-atom MD simulations.

Lennard-Jones potential between the oxygen sites on two SPC/E water molecules are  $A_{oo} = 2.6334 \times 10^8 \text{ Da } \text{Å}^{14} \text{ ps}^{-2}$  and  $B_{oo} = 2.6171 \times 10^5 \text{ Da } \text{Å}^8 \text{ ps}^{-2}$ .

We consider a cube of side  $L = 24.83 \text{ Å}$  containing 511 water molecules and 1 ion, i.e. we have  $8 \times 3 = 512$  molecules in our simulation box. In the following section, we use standard NVT simulations where the temperature is controlled using Nosé-Hoover thermostat [20, 21] and the number of particles is kept constant by implementing periodic boundary conditions. In particular, we assume that our simulation box is surrounded by periodic copies of itself. Then the long-range (Coulombic) interactions can be computed using several different approaches, including the Ewald summation or the reaction field method [22, 23]. We use the cutoff sphere of radius  $L/2$  and the reaction field correction as implemented in [1]. This approach is more suitable for multiscale methods (studied later in Section 6) than the Ewald summation technique. The MD timestep is for all MD simulations in this paper chosen as  $\Delta t = 10^{-3} \text{ ps} = 1 \text{ fs}$ .

### 3. Parametrization of the coarse-grained model of ion

In MD simulations, an ion is described by its position  $X \equiv [X_1, X_2, X_3]$  and velocity  $V \equiv [V_1, V_2, V_3]$  which evolve according to

$$dX_i = V_i dt, \tag{3.1}$$

$$M dV_i = F_i dt, \quad i = 1, 2, 3, \tag{3.2}$$

where  $M$  is the mass of the ion (given in Table 1) and  $F \equiv [F_1, F_2, F_3]$  is the force acting on the ion. We use all-atom MD simulations as described in Section 2 to estimate diffusion coefficient  $D$  and second moments of  $V_i$  and  $U_i = F_i/M$ ,  $i = 1, 2, 3$ . They are given for four ions considered in Table 2. To estimate  $\langle U_i^2 \rangle$ , we calculate the average force in the  $i$ -th direction  $\langle F_i^2 \rangle$  where  $\langle \cdot \rangle$  denotes an average over sufficiently large time interval (nanosecond) of MD simulations. Taking into account the symmetry of the problem, we estimate  $\langle U_i^2 \rangle = \langle F_i^2 \rangle / M^2$  as the average over all three dimensions

$$\frac{\langle U_1^2 \rangle + \langle U_2^2 \rangle + \langle U_3^2 \rangle}{3}$$

This value is reported in Table 2. In the same way, the reported values of  $\langle V_i^2 \rangle$  are computed as averages over all three dimensions. Diffusion constant  $D$  can be

ion	$\eta_1$ [ps <sup>-2</sup> ]	$\eta_2$ [ps <sup>-1</sup> ]	$\eta_3$ [ps <sup>-2</sup> ]	$\eta_4$ [Å ps <sup>-7/2</sup> ]
K <sup>+</sup>	768.7	152.5	$3.393 \times 10^3$	$7.094 \times 10^4$
Na <sup>+</sup>	$2.044 \times 10^3$	166.1	$4.020 \times 10^3$	$1.717 \times 10^5$
Ca <sup>2+</sup>	$3.026 \times 10^3$	190.2	$4.933 \times 10^3$	$1.874 \times 10^5$
Cl <sup>-</sup>	940.0	189.7	$4.524 \times 10^3$	$1.061 \times 10^5$

Table 3. Values of  $\eta_j$ ,  $j = 1, 2, 3, 4$ , estimated using all-atom MD simulations.

estimated by calculating mean square displacements or velocity autocorrelation functions. In Table 2, we report the values of D which were estimated in [1] by calculating mean square displacements. Let us consider the coarse-grained model (1.2)–(1.5) and let  $\langle \cdot \rangle$  denotes an average over many realizations of a stochastic process. Multiplying equations (1.3) and (1.4) by  $V_i$  and  $U_i$ , respectively, we obtain the following ODEs for second moments:

$$\frac{d}{dt} \langle V_i^2 \rangle = 2 \langle U_i V_i \rangle, \tag{3.3}$$

$$\frac{d}{dt} \langle U_i^2 \rangle = -2\eta_1 \langle U_i V_i \rangle + 2 \langle U_i Z_i \rangle. \tag{3.4}$$

Consequently, we obtain that  $\langle U_i V_i \rangle = 0$  and  $\langle U_i Z_i \rangle = 0$  at steady state. Multiplying equations (1.3)–(1.5) by  $V_i$ ,  $U_i$  and  $Z_i$ , and taking averages, we obtain

$$\frac{d}{dt} \langle U_i V_i \rangle = \langle U_i^2 \rangle - \eta_1 \langle V_i^2 \rangle + \langle V_i Z_i \rangle, \tag{3.5}$$

$$\frac{d}{dt} \langle V_i Z_i \rangle = \langle U_i Z_i \rangle - \eta_2 \langle V_i Z_i \rangle - \eta_3 \langle U_i V_i \rangle. \tag{3.6}$$

Using  $\langle U_i V_i \rangle = 0$  and  $\langle U_i Z_i \rangle = 0$ , we obtain that  $\langle V_i Z_i \rangle = 0$  at steady state and

$$\eta_1 = \frac{\langle U_i^2 \rangle}{\langle V_i^2 \rangle}. \tag{3.7}$$

This equation is used in Table 3 to estimate  $\eta_1$  using the MD averages  $\langle U_i^2 \rangle$  and  $\langle V_i^2 \rangle$  which are given in Table 2. Since we know the value of  $\eta_1$ , we can also estimate the value of  $\langle Z_i^2 \rangle$  by calculating the second moment of

$$\langle Z_i^2 \rangle \approx \left\langle \left( \frac{U_i(t + \Delta t) - U_i(t)}{\Delta t} + \eta_1 V_i \right)^2 \right\rangle. \tag{3.8}$$

This value is reported in the last column of Table 2. Multiplying equation (1.4) by  $Z_i$  and equation (1.5) by  $U_i$ , we obtain

$$\frac{d}{dt} \langle U_i Z_i \rangle = \langle Z_i^2 \rangle - \eta_1 \langle V_i Z_i \rangle - \eta_2 \langle U_i Z_i \rangle - \eta_3 \langle U_i^2 \rangle. \tag{3.9}$$

Using  $\langle U_i Z_i \rangle = 0$  and  $\langle V_i Z_i \rangle = 0$ , we obtain at steady state

$$\frac{d}{dt}\langle U_i Z_i \rangle = \langle Z_i^2 \rangle - \eta_1 \langle V_i Z_i \rangle - \eta_2 \langle U_i Z_i \rangle - \eta_3 \langle U_i^2 \rangle. \tag{3.9}$$

Multiplying equation (1.2) by  $X_i$ ,  $V_i$ ,  $U_i$  and  $Z_i$  and equations (1.3)–(1.5) by  $X_i$  and taking averages, we obtain the following system of ODEs for second moments:

$$\frac{d}{dt}\langle X_i^2 \rangle = 2\langle X_i V_i \rangle, \tag{3.11}$$

$$\frac{d}{dt}\langle X_i V_i \rangle = \langle V_i^2 \rangle + \langle X_i U_i \rangle, \tag{3.12}$$

$$\frac{d}{dt}\langle X_i U_i \rangle = \langle U_i V_i \rangle - \eta_1 \langle X_i V_i \rangle + \langle X_i Z_i \rangle, \tag{3.13}$$

$$\frac{d}{dt}\langle X_i Z_i \rangle = \langle V_i Z_i \rangle - \eta_2 \langle X_i Z_i \rangle - \eta_3 \langle X_i U_i \rangle. \tag{3.14}$$

Consequently, we obtain at steady state  $\langle X_i V_i \rangle = D$ ,  $\langle X_i U_i \rangle = -\langle V_i^2 \rangle / \eta_1$  and  $\langle X_i Z_i \rangle = \langle V_i Z_i \rangle / \eta_1$

$$\eta_2 = -\frac{\eta_3 \langle X_i U_i \rangle}{\langle X_i Z_i \rangle} = \frac{\eta_3 \langle V_i^2 \rangle}{\eta_1 D}.$$

Using (3.7) and (3.10), we have

$$\eta_2 = \frac{\langle Z_i^2 \rangle}{D} \left( \frac{\langle V_i^2 \rangle}{\langle U_i^2 \rangle} \right)^2. \tag{3.15}$$

Finally, multiplying equation (1.5) by  $Z_i$ , we obtain

$$\frac{d}{dt}\langle Z_i^2 \rangle = -2\eta_2 \langle Z_i^2 \rangle - 2\eta_3 \langle U_i Z_i \rangle + \eta_4^2. \tag{3.16}$$

Consequently, we obtain at steady state

$$\eta_4^2 = 2\eta_2 \langle Z_i^2 \rangle.$$

Using (3.15), we get

$$\eta_4 = \sqrt{\frac{2 \langle V_i^2 \rangle \langle Z_i^2 \rangle}{D \langle U_i^2 \rangle}}. \tag{3.17}$$

The values calculated by (3.10), (3.15) and (3.17) are presented in Table 3.

#### 4. Accuracy of the coarse-grained model of ion

The coarse-grained model (1.2)–(1.5) has four parameters  $\eta_i$ ,  $i = 1, 2, 3, 4$ . To parameterize this model, we have used four quantities estimated from detailed MD simulations, diffusion constant  $D$  and steady state values of  $\langle X_i V_i \rangle$ ,  $\langle X_i U_i \rangle$  and  $\langle X_i Z_i \rangle$ . In particular, the coarse-grained model (1.2)–(1.5) will give the same values of these four quantities, including the value of diffusion constant  $D$  which is the sole parameter of the BD model (1.1). In this section, we explain why the coarsegrained description given by (1.2)–(1.5) can be used as an intermediate model to couple BD and MD models.

We begin by illustrating why Langevin dynamics (which is used in [12] for a similar multiscale problem) is not suitable for all-atom MD simulations studied in this paper. In [12], a few (heavy) particles with mass  $M$  and radius  $R$  are considered in the heat bath consisting of a large number of light point particles with masses  $m \ll M$ . The collisions of particles are without friction, which means that post-collision velocities can be computed using the conservation of momentum and energy. In this case, it can be shown that the description of heavy particles converges in an appropriate limit to Brownian motion given by equation (1.1). One can also show that the model converges to Langevin dynamics (in the limit  $m/M \rightarrow 0$ ) [24, 25, 26]:

$$dX_i = V_i dt, \tag{4.1}$$

$$dV_i = -\gamma V_i dt + \gamma\sqrt{2D} dW_i, \quad i = 1, 2, 3, \tag{4.2}$$

where  $X \equiv [X_1, X_2, X_3]$  is the position of a diffusing molecule,  $V \equiv [V_1, V_2, V_3]$  is its velocity,  $D$  is the diffusion coefficient and  $\gamma$  is the friction coefficient. In [12], Langevin dynamics (4.1)–(4.2) is used as an intermediate model which enables the implementation of BD description (1.1) and the original detailed model in different parts of the computational domain.

Langevin dynamics (4.1)–(4.2) describes a diffusing particle in terms of its position and velocity, i.e. it uses the same independent variables for the description of an ion as the MD model (3.1)–(3.2). Langevin dynamics can be further reduced to BD model (1.1) in the overdamped limit  $\gamma \rightarrow \infty$ . However, it cannot be used as an intermediate model between BD and all-atom MD simulations considered in this paper, because it does not correctly describe the ion behaviour at times comparable to the MD timestep  $\Delta t$ . To illustrate this, let us parameterize Langevin dynamics (4.1)–(4.2) using diffusion constant  $D$  and the second velocity moment  $\langle V_i^2 \rangle$  estimated from all-atom MD simulations. To get the same second moment of velocity, Langevin dynamics requires that we choose

$$\gamma = \frac{\langle V_i^2 \rangle}{D}. \tag{4.3}$$

Discretizing equation (4.2), the ion acceleration during one time step is

$$\frac{V_i(t + \Delta t) - V_i(t)}{\Delta t} = -\gamma V_i(t) + \gamma\sqrt{\frac{2D}{\Delta t}} \xi_i \tag{4.4}$$

where  $[\xi_1, \xi_2, \xi_3]$  is a vector of normally distributed random numbers with zero mean and unit variance. Using (4.3), the second moment of the right hand side of (4.4) is

$$\gamma^2 \left( \langle V_i^2 \rangle + \frac{2D}{\Delta t} \right) = \frac{\langle (V_i^2) \rangle^2}{D^2} + \frac{2\langle (V_i^2) \rangle}{D \Delta t}. \tag{4.5}$$

Using the MD values of  $D$  and  $\langle V_i^2 \rangle$  for  $K^+$  which are given in Table 2 and using MD timestep  $\Delta t = 10^{-3}$  ps, we obtain that the second moment (4.5) is equal to  $4.44 \times 10^5 \text{ \AA}^2 \text{ ps}^{-4}$ . On the other hand,  $\langle V_i^2 \rangle$  estimated from all-atom

MD simulations and given in Table 2 is  $4.86 \times 10^3 \text{ \AA}^2 \text{ ps}^{-4}$  which is one hundred times smaller. The main reason for this discrepancy is that Langevin dynamics postulates that the random force in equation (4.2) acting on the particle at time  $t$  is not correlated to the random force acting on the particle at time  $t + \Delta t$ . However, this is not true for all-atom MD simulations where random force terms at subsequent time steps are highly correlated.

Since Langevin dynamics is not suitable for coupling MD and BD models, we need to introduce a stochastic model of ion behaviour which is more complicated than Langevin dynamics. The coarse-grained model (1.2)–(1.5) studied in this paper is a relatively simple example of such a model. Its parametrization, discussed in Section 3, guarantees that the coarse-grained model (1.2)–(1.5) well approximates all-atom MD simulations at steady state. They both have the same value of diffusion constant  $D$  and steady state values of  $\langle hV_i^2 \rangle$ ,  $\langle hU_i^2 \rangle$  and  $\langle hZ_i^2 \rangle$ . Next, we show that the coarse-grained model (1.2)–(1.5) also compares well with all MD simulations at shorter timescales. We consider the rate of change of acceleration (jerk or the scaled derivative of force). We define the average jerk as a function of current velocity and acceleration of the ion:

$$J(v, u) = \lim_{\tau \rightarrow 0} \frac{\langle U_i(t + \tau) - u \mid V_i(t) = v, U_i(t) = u \rangle}{\tau} \tag{4.6}$$

To estimate  $J(v, u)$  from all-atom MD simulations, we calculate the rate of change of acceleration during each MD time step

$$J(v, u) \approx \frac{\langle U_i(t + \Delta t) - u \mid V_i(t) = v, U_i(t) = u \rangle}{\Delta t} \tag{4.7}$$

i.e. we run a long (nanosecond) MD simulation, calculate the values of  $(U_i(t + \Delta t) - U_i(t))/\Delta t$  during every time step and record their average in two-variable array  $J(v, u)$  indexed by binned values of  $V_i(t) = v$  and  $U_i(t) = u$ . Since the estimated  $J(v, u)$  only weakly depends on  $u$ , we visualize our results in Figure 1 using two functions of one variable,  $v$ , namely

$$J_1(v) = J(v, 0), \quad \text{and} \quad J_2(v) = \int_{-\infty}^{\infty} J(v, u) p_u(u) du, \tag{4.8}$$

where  $p_u(u)$  is the steady state distribution of  $U_i$  estimated from the same long time MD trajectory. As before, we use all three dimensions to calculate the averages  $J(v, u)$  and  $p_u(u)$ . Function  $J_1(v)$  (which gives jerk at the most likely value of  $U_i$ ) is plotted using crosses and function  $J_2(v)$ , the average over  $U_i$  variable, is plotted using circles in Figure 1. In order to compare all-atom MD simulations with the coarse-grained model (1.2)–(1.5), we calculate the corresponding jerk matrix  $J(v, u)$  for the coarse-grained model. We denote by  $p(v, u, z)$  the stationary distribution of the stochastic process (1.3)–(1.5), i.e.  $p(v, u, z) dv du dz$  is the probability that  $V_i(t) \in [v, v + dv)$ ,  $U_i(t) \in [u, u + du)$  and  $Z_i(t) \in [z, z + dz)$ . Then the jerk matrix (4.6) of the coarse-grained model (1.2)–(1.5) is

$$J(v, u) = \int_{-\infty}^{\infty} \lim_{\tau \rightarrow 0} \frac{\langle U_i(t + \tau) - u \mid V_i(t) = v, U_i(t) = u, Z_i(t) = z \rangle}{\tau} p(v, u, z) dz.$$



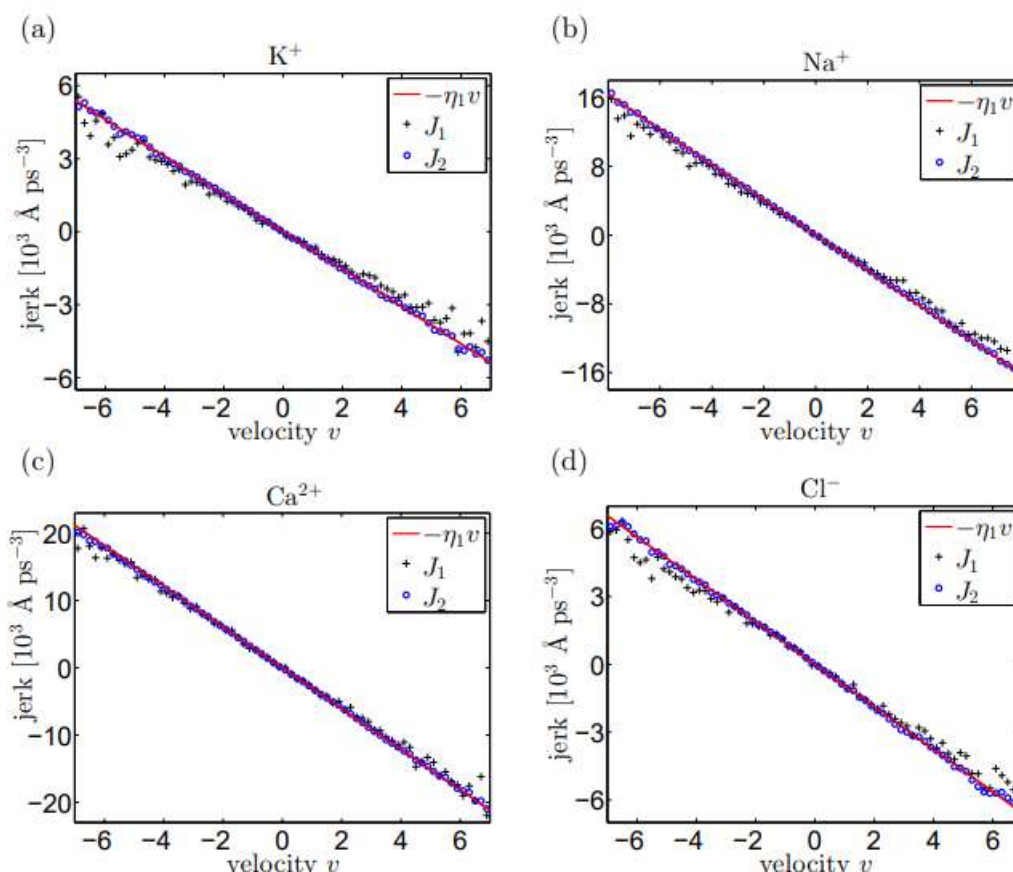


Figure 1. Comparison of the rate of change of acceleration (jerk) computed by all-atom MD simulations and by the coarse-grained model (1.2)–(1.5). MD results are visualized using functions  $J_1(v)$  (black crosses) and  $J_2(v)$  (blue circles) defined by equation (4.8). The result obtained by the coarse-grained model is given by formula (4.10) (red solid line). We consider (a)  $K^+$  ion; (b)  $Na^+$  ion; (c)  $Ca^{2+}$  ion and (d)  $Cl^-$  ion. Parameters are given in Tables 1 and 3.

Using (1.4), we rewrite it as

$$J(v, u) = \int_{-\infty}^{\infty} (-\eta_1 v + z) p(v, u, z) dz. \tag{4.9}$$

The stationary distribution  $p(v, u, z)$  of (1.3)–(1.5) is Gaussian with mean  $[0, 0, 0]^T$  and stationary covariance matrix:

$$\frac{1}{2\eta_1\eta_2\eta_3} \begin{pmatrix} \eta_4^2 & 0 & 0 \\ 0 & \eta_1\eta_4^2 & 0 \\ 0 & 0 & \eta_1\eta_3\eta_4^2 \end{pmatrix}.$$

Consequently, equation (4.9) implies

$$J(v, u) = -\eta_1 v. \tag{4.10}$$

ion	$\lambda_1$ [ps <sup>-1</sup> ]	$\lambda_2$ [ps <sup>-1</sup> ]	$\lambda_3$ [ps <sup>-1</sup> ]	$t_1^*$ [ps]	$t_2^*$ [ps]
K <sup>+</sup>	-127.0	-12.75 + 27.58 i	-12.75 - 27.58 i	$3.08 \times 10^{-2}$	$-9.39 \times 10^{-3}$
Na <sup>+</sup>	-140.1	-12.99 + 47.47 i	-12.99 - 47.47 i	$6.15 \times 10^{-3}$	$-2.35 \times 10^{-2}$
Ca <sup>2+</sup>	-163.1	-13.58 + 57.84 i	-13.58 - 57.84 i	$1.47 \times 10^{-3}$	$-2.48 \times 10^{-2}$
Cl <sup>-</sup>	-162.9	-13.41 + 30.25 i	-13.41 - 30.25 i	$2.50 \times 10^{-2}$	$-1.13 \times 10^{-2}$

Table 4. Eigenvalues  $\lambda_j, j = 1, 2, 3$ , of matrix  $B$  defined by (5.2) and time shifts  $t_1^*$  and  $t_2^*$ . Symbol  $i$  denotes the imaginary unit.

In Figure 1, we plot (4.10) using the red solid line. The comparison with all atom MD results (circles and squares) is excellent for all four ions considered in this paper. In particular, we have shown that the coarse-grained model (1.2)–(1.5) provides a good description of the rate of change of acceleration (jerk) at the MD timescale. We make use of this property in Section 6 where we use the same time step ( $\Delta t = 10^{-3}$  ps) for both the coarse-grained model (1.2)–(1.5) and all-atom MD simulations. The coarse-grained model (1.2)–(1.5) can also be coupled with BD description (1.1), which uses much larger time steps, as we show in the next section.

**5. From the coarse-grained model (1.2)–(1.5) to Brownian dynamics**

Let us consider the three-variable subsystem (1.3)–(1.5) of the coarse-grained model. Denoting  $y_i = [V_i, U_i, Z_i]$ , equations (1.3)–(1.5) can be written in vector notation as follows

$$dy_i = B y_i dt + b dW_i, \tag{5.1}$$

where matrix  $B \in \mathbb{R}^{3 \times 3}$  and vector  $b \in \mathbb{R}^3$  are given as

$$B = \begin{pmatrix} 0 & 1 & 0 \\ -\eta_1 & 0 & 1 \\ 0 & -\eta_3 & -\eta_2 \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} 0 \\ 0 \\ \eta_4 \end{pmatrix}. \tag{5.2}$$

Let us denote the eigenvalues and eigenvectors of  $B$  as  $\lambda_j$  and  $v_j = [v_{1j}, v_{2j}, v_{3j}]$ ,  $j = 1, 2, 3$ , respectively. The eigenvalues of  $B$  are the solutions of the characteristic polynomial

$$\lambda^3 + \eta_2 \lambda^2 + (\eta_1 + \eta_3)\lambda + \eta_1 \eta_2 = 0.$$

Since  $\eta_1, \eta_2$  and  $\eta_3$  are positive parameters, we conclude that real parts of all three eigenvalues are negative and lie in interval  $(-\eta_2, 0)$ . Using the values of  $\eta_j, j = 1, 2, 3$ , given in Table 3, we present the values of eigenvalues  $\lambda_j, j = 1, 2, 3$ , in Table 4. The eigenvalues  $\lambda_j, j = 1, 2, 3$ , are distinct. The general solution of the

SDE system (5.1) can be written as follows [27]

$$y_i(t) = \Phi(t) c + \Phi(t) \int_0^t \Phi^{-1}(s) b dW_i, \tag{5.3}$$

where  $c \in \mathbb{R}^3$  is a constant vector determined by initial conditions and matrix  $\Phi(t) \in \mathbb{R}^{3 \times 3}$  is given as  $\Phi(t) = [\exp(\lambda_1 t)v_1 \mid \exp(\lambda_2 t)v_2 \mid \exp(\lambda_3 t)v_3]$ , i.e. each column is a solution of the

ODE system  $dy_i = B y_i dt$ . Considering deterministic initial conditions, equation (5.3) implies that the process is Gaussian at any time  $t > 0$ . Equations for means, variances and covariances then uniquely determine the distribution of  $y_i(t)$  for  $t > 0$ . Equations for means can be written in the vector form as  $dh_{y_i} = B h_{y_i} dt$ . Equations for variances and covariances are given in Section 3 as equations (3.3)–(3.6), (3.9), (3.11)–(3.14) and (3.16).

There are two important conclusions of the above analysis. First of all, eigenvalues  $\lambda_j$ ,  $j = 1, 2, 3$ , given in Table 4 satisfy

$$\lambda_1 < \text{Re } \lambda_2 = \text{Re } \lambda_3 < 0,$$

where  $\text{Re}$  denotes the real part of a complex number. There is a spectral gap between the first eigenvalue and the complex conjugate pair of eigenvalues. If we used this spectral gap, we could reduce the system to two evolution equations for times  $t \gg 1/|\lambda_1|$ . However, there is no spectral gap to reduce the system to Langevin dynamics (4.1)–(4.2). In particular, we again confirm our conclusion that a coarse-grained approximation of ion behaviour is not given in terms of Langevin dynamics. Our second conclusion is that on a picosecond time scale, we can assume stationarity in (5.1) to get

$$dX_i = \frac{\eta_4}{\eta_1 \eta_2} dW_i, \quad i = 1, 2, 3. \tag{5.4}$$

Using (3.7), (3.15) and (3.17), we have

$$\frac{\eta_4}{\eta_1 \eta_2} = \sqrt{2D}.$$

Consequently, equation (5.4) is equivalent to BD description (1.1). The convergence of (1.2)–(1.5) to the BD model is illustrated in Figure 2(a). We solve the system of 10 ODEs for variances and covariances given as equations (3.3)–(3.6), (3.9), (3.11)–(3.14) and (3.16). We consider (deterministic) zero initial conditions, i.e.  $X_i(0) = V_i(0) = U_i(0) = Z_i(0) = 0$ . All moments are then initially equal to zero. We plot the mean square displacement  $\langle X_i^2 \rangle$  as a function of time. We compare it with the mean square displacement of BD model (1.1) which is given as  $2Dt$ . We observe that there is an approximately constant shift, denoted  $t^* - 1$ , between both solutions for times  $t > 0.2$  ps. We illustrate this further by plotting  $\langle X_i^2 \rangle(t + t^* - 1)$  in Figure 2(a). The values of shift  $t^* - 1$  for different ions estimated by solving the ODEs for second moments with zero initial conditions are given in Table 4.

Next, we show how the BD model (1.1) and the coarse-grained model (1.2)–(1.5) can be used in different parts of the computational domain. This coupling will form one component of multiscale methodology developed in Section 6. BD algorithms based on equation (1.1) have been implemented in a number of methods designed for spatio-temporal modelling of intracellular processes,

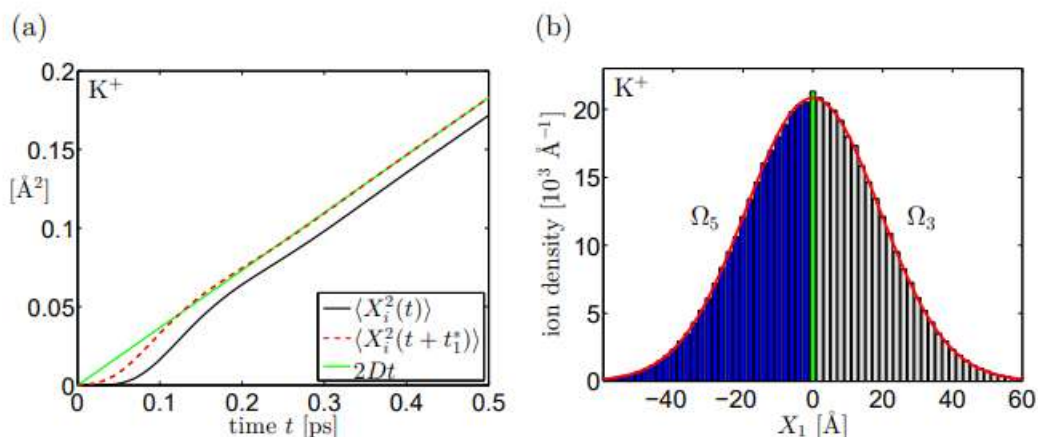


Figure 2. (a) Comparison of the coarse-grained model (1.2)–(1.5) and BD description (1.1) for  $K^+$  ion. The mean square displacement computed by solving 10 ODEs (3.3)–(3.6), (3.9), (3.11)– (3.14) and (3.16) with zero initial conditions (black solid line). The same curve shifted by the value of  $t * 1$  is plotted as a red dashed line. (b) Test of accuracy of the multiscale approach in  $\Omega_3 \cup \Omega_4 \cup \Omega_5$  for  $K^+$  ion. Histogram obtained by simulating 106 ions over time 103 ps is compared with the analytical result (5.6) (red solid line). Grey bars show the ion density in  $\Omega_3$ , the green bar shows the ion density in  $\Omega_4$  and blue bars show the ion density in  $\Omega_5$ . Parameters are given in Tables 1 and 3.

including Smoldyn [28], MCell [29] and Green’s-function reaction dynamics [30]. Smoldyn discretizes (1.1) using a fixed BD time step  $\Delta T$ , i.e. it computes the time evolution of the position  $X \equiv X(t)$  of each molecule by

$$X_i(t + \Delta t) = X_i(t) + \sqrt{2D\Delta T} \xi_i, \quad i = 1, 2, 3, \quad (5.5)$$

where  $[\xi_1, \xi_2, \xi_3]$  is a vector of normally distributed random numbers with zero mean and unit variance. We use discretization (5.5) of BD model (1.1) in this paper. BD time step  $\Delta T$  has to be chosen much larger than the MD time step  $\Delta t$ . We use  $\Delta T = 0.5$  ps, but any larger time step would also work well. We could also use a variable time step, as implemented in the Green’s Function Reaction Dynamics [30].

In Section 6, we consider all-atom MD simulations in domain  $\Omega \subset R^3$ . Our main goal is to design a multiscale approach which can compute spatio-temporal statistics with the MD-level of detail in relatively small subdomain  $\Omega_1 \subset \Omega$  by using BD model (5.5) in the most of the rest of the computational domain. This is achieved by decomposing domain  $\Omega$  into five subdomains  $\Omega_j, j = 1, 2, 3, 4, 5$  (see equation (6.1) and discussion in Section 6). We use MD in  $\Omega_1$ , the coarsegrained model (1.2)–(1.5) in  $\Omega_3$  and the BD model (5.5) in  $\Omega_5$ . The remaining two subdomains,  $\Omega_2$  and  $\Omega_4$ , are two overlap (hand-shaking) regions where two different simulation approaches can be used at the same time [12, 31]. In the rest of this section, we focus on simulations in region  $\Omega_3 \cup \Omega_4 \cup \Omega_5$  which concerns coupling the coarse-grained model (1.2)–(1.5) with the BD model (5.5). We use the coarse-grained model in  $\Omega_3 \cup \Omega_4$  and the BD model (5.5) in  $\Omega_4 \cup \Omega_5$ . In particular, we use both models in the overlap region  $\Omega_4$ . Each particle which is initially in  $\Omega_3$  is simulated according to (1.2)–(1.5) (discretized using time step  $\Delta t$ ) until it

enters  $\Omega_5$ . Then we use (5.5) to evolve the position of a particle (over BD time steps of length  $\Delta T$ ) until it again enters  $\Omega_3$  when we switch the description back from the BD model to the coarse-grained model. In order to do this, we have to initialize variables  $V_i$ ,  $U_i$  and  $Z_i$ ,  $i = 1, 2, 3$ . We use deterministic initial conditions,  $V_i(0) = U_i(0) = Z_i(0) = 0$ , discussed above.

In Figure 2(b), we present an illustrative simulation where  $\Omega_3 \cup \Omega_4 \cup \Omega_5 = \mathbb{R}^3$  for simplicity. We use  $\Omega_3 = (h, \infty) \times \mathbb{R}^2$ ,  $\Omega_4 = [-h, h] \times \mathbb{R}^2$  and  $\Omega_5 = (-\infty, -h) \times \mathbb{R}^2$ , where  $h = 1 \text{ \AA}$ . We report averages over 106 simulations of ions, half of them are initiated at  $X(0) = [h, 0, 0]$ , i.e. they initially follow the coarse-grained model (1.2)–(1.5) with zero initial condition for other variables ( $V_i(0) = U_i(0) = Z_i(0) = 0$ ). The second half of ions are initiated at  $X(0) = [-h, 0, 0]$ , i.e. they initially follow BD description (5.5). We plot the (marginal) distribution of ions along the first coordinate ( $X_1$ ) at time 103 ps in Figure 2(b). The computed histogram is plotted using bins of length  $2 \text{ \AA}$ , i.e. the overlap region  $\Omega_4$  is equal to one bin (visualized as a green bar). Grey (resp. blue) bars show the density of ions in  $\Omega_3$  (resp.  $\Omega_5$ ). We compare our results with the analytical distribution computed for BD description (1.1) at time  $t = 103 \text{ ps}$  given by

$$\rho(x_1) = \frac{10^6}{4\sqrt{\pi Dt}} \left( \exp \left[ -\frac{(x_1 - h)^2}{4Dt} \right] + \exp \left[ -\frac{(x_1 + h)^2}{4Dt} \right] \right). \quad (5.6)$$

The computed histogram compares well with (5.6), although we can observe a small error: the green bar is slightly taller than the corresponding value of (5.6). If we wanted to further improve the accuracy, we could take into account that there is time shift  $t^* 1$ , discussed above, introduced to the multiscale approach by using the deterministic initial conditions,  $V_i(0) = U_i(0) = Z_i(0) = 0$ , for ions entering domain  $\Omega_3$ . Another possibility is to sample the initial condition for  $V_i$ ,  $U_i$  and  $Z_i$  from a suitable distribution. If we use the stationary distribution of subsystem (1.3)–(1.5), then  $\langle V_i^2 \rangle$  does not evolve and is equal to

$$\langle V_i^2 \rangle = \frac{\eta_4^2}{2\eta_1\eta_2\eta_3}.$$

Substituting this constant for  $\langle V_i^2 \rangle$  into (3.12), the system of 10 ODEs for second moments of (1.2)–(1.5) simplifies to 4 ODEs (3.11)–(3.14). Solving system (3.11)–(3.14) with zero initial conditions (assuming  $X_i(0) = 0$ ), we can again compute the mean square displacement. As in Figure 2(a), it can be shifted in time to better match with the BD result,  $2Dt$ . We denote this time shift as  $t^* 2$ . Its values are given in Table 4. We observe that  $t^* 2$  is negative and  $t^* 1$  is positive for all four ions considered in Table 4. Both time shifts  $t^* 1$  and  $t^* 2$  (together with optimizing size  $h$  of the overlap region) could be used to further improve the accuracy of multiscale simulations in  $\Omega_3 \cup \Omega_4 \cup \Omega_5$  [12]. However, our main goal is to introduce a multiscale approach which can use all-atom MD simulations in  $\Omega_1$ . Since MD simulations are computationally intensive, we will only consider 100 realizations of the multiscale method in Section 6. In particular, the Monte Carlo error will be larger than the error observed in Figure 2(b). Thus, we can use the above approach in  $\Omega_3 \cup \Omega_4 \cup \Omega_5$  without introducing observable errors in the multiscale method developed in the next section.

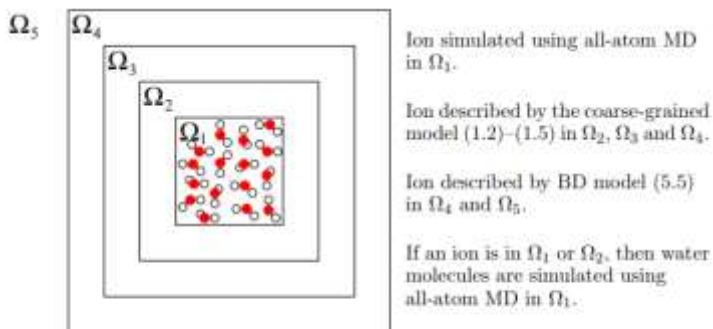


Figure 3. Schematic of multiscale set up. Note that the schematic is drawn in two spatial dimensions to enable better visualization, but all models are formulated and simulated in three spatial dimensions.

### 6. Coupling all-atom MD and BD

Let us consider all-atom MD in domain  $\Omega \subset \mathbb{R}^3$  which is so large that direct MD simulations would be too computationally expensive. Let us assume that a modeller only needs to consider the MD-level of detail in a relatively small subdomain  $\Omega_1 \subset \Omega$  while, in the rest of the computational domain, ions are transported by diffusion and BD description (1.1) is applicable. For example, domain  $\Omega_1$  could include binding sites for ions or (parts of) ion channels. In this paper, we do not focus on a specific application. Our goal is to show that the coarse-grained model (1.2)–(1.5) is an intermediate model between all-atom MD and BD which enables the use of both methods during the same dynamic simulation. To achieve this, we decompose domain  $\Omega$  into five subdomains, denoted  $\Omega_j, j = 1, 2, 3, 4, 5$ , see Figure 3. These sets are considered pairwise disjoint (i.e.  $\Omega_i \cap \Omega_j = \emptyset$  for  $i \neq j$ ) and

$$\Omega = \Omega_1 \cup \Omega_2 \cup \Omega_3 \cup \Omega_4 \cup \Omega_5. \tag{6.1}$$

In our illustrative simulations, we consider the behaviour of one ion. If the ion is in subdomain  $\Omega_1$ , then we use all-atom MD simulations as described in Section 2. In particular, the force between the ion and a water molecule is obtained by differentiating potential (2.1), provided that the distance between the ion and the water molecule is less than the cutoff distance ( $L/2$ ). Let us denote the force exerted by the ion on the water molecule by  $F_{iw}(r_{i0}, r_{i1}, r_{i2})$ , where  $r_{i0}$  (resp.,  $r_{i1}$  and  $r_{i2}$ ) is the distance between the ion and the oxygen site (resp., the first and second hydrogen sites) on the water molecule. We use periodic boundary conditions for water molecules in  $\Omega_1$ .

Whenever the ion leaves  $\Omega_1$ , it enters  $\Omega_2$  where we simulate its behaviour using the coarse-grained model (1.2)–(1.5). We still simulate water molecules in  $\Omega_1$  and we allow them to experience additional forces exerted by the ion which is present

in  $\Omega_2$ . These forces have the same functional form,  $F_{iw}$ , as in MD, but they have modified arguments as follows

$$\tilde{F}_{iw}(r_{i0} + \omega \text{dist}(\mathbf{X}, \Omega_1), r_{i1} + \omega \text{dist}(\mathbf{X}, \Omega_1), r_{i2} + \omega \text{dist}(\mathbf{X}, \Omega_1)), \tag{6.2}$$

where  $\omega \geq 0$  is a parameter and  $\text{dist}(X, \Omega_1)$  is the (closest) distance between the ion at position  $X$  and subdomain  $\Omega_1$ . If the ion is in region  $\Omega_3 \cup \Omega_4 \cup \Omega_5$ , then water molecules in  $\Omega_1$  are no longer simulated. We use the coarse-grained model (1.2)–(1.5) to simulate the ion

behaviour in  $\Omega_3$  and the BD model (5.5) in  $\Omega_5$ . Overlap region  $\Omega_4$  is used to couple these simulation methods as explained in Section 5

In Section 5, we have already presented illustrative simulations to validate the multiscale modelling strategy chosen in region  $\Omega_3 \cup \Omega_4 \cup \Omega_5$ . Next, we focus on testing and explaining the multiscale approach chosen to couple region  $\Omega_1$  with  $\Omega_2$ . The key idea is given by force term (6.2) which is used for MD simulations of water molecules in  $\Omega_1$  when an ion is in  $\Omega_2$ . This force term has two important properties:

- (i) If an ion is on the boundary of  $\Omega_1$ , i.e.  $X \in \partial\Omega_1$ , then  $\text{dist}(X, \Omega_1) = 0$  and force (6.2) is equal to force term  $F_{iw}(r_{i0}, r_{i1}, r_{i2})$  used in  $\Omega_1$ .
- (ii) If  $\omega \text{dist}(X, \Omega_1) \geq L/2$ , then force (6.2) is equal to zero.

Property (i) implies that formula (6.2) continuously extends the force term used in MD. In particular, water molecules do not experience abrupt changes of forces when the ion crosses boundary  $\partial\Omega_1$ . Property (ii) is a consequence of the cutoff distance used (together with the reaction field correction [1]) to treat long-range interactions. In our illustrative simulations, we use

$$\Omega_1 = \left[-\frac{L}{2}, \frac{L}{2}\right]^3 \quad \text{and} \quad \Omega_2 = \left[-\frac{L}{2} - \frac{L}{2\omega}, \frac{L}{2} + \frac{L}{2\omega}\right]^3 \setminus \Omega_1. \quad (6.3)$$

Property (ii) implies that extra force (6.2) is equal to zero on boundary  $\partial\Omega_2 \setminus \partial\Omega_1$  which is the boundary between regions  $\Omega_2$  and  $\Omega_3$ . This is consistent with the assumption that ions in region  $\Omega_3 \cup \Omega_4 \cup \Omega_5$  do not interact with water molecules in region  $\Omega_1$ .

If an ion is in  $\Omega_1$ , we use all-atom MD as formulated in Section 2. Periodic boundary conditions are implemented in MD simulations. Water molecules are subject to forces exerted by the ion at its real position  $X$  in  $\Omega_1$ , but also by its copies at periodic locations  $X + (iL, jL, kL)$  where  $i, j, k \in \mathbb{Z}$ . When the ion moves to  $\Omega_2$ , one of its copies is in  $\Omega_1$ . Force term (6.2) is designed in such a way, that the strength of interaction decreases (for every copy of the ion) with the distance,  $\text{dist}(X, \Omega_1)$ , between the real position of the ion and  $\Omega_1$ . In particular, force term (6.2) ensures that there are continuous changes of all forces when the ion moves between regions  $\Omega_1, \Omega_2$  and  $\Omega_3$

In Figure 4, we present results of simulations of  $K^+$  ion in region  $\Omega_1 \cup \Omega_2$ . We consider 100 realizations of a multiscale simulation with one ion. Its initial position is  $X(0) = [-L/2, 0, 0]$  which lies on boundary  $\partial\Omega_1$ . We simulate each realization for time 10 ps which is short enough that all trajectories stay inside the ball of radius  $L/2$  centred at  $X(0)$ . Then  $X_1$ -coordinate of the trajectory determines whether the ion is in  $\Omega_1$  or  $\Omega_2$ . If  $X_1(t) \geq -L/2$ , then the ion is in  $\Omega_1$  and it is simulated using all-atom MD. If  $X_1(t) < -L/2$ , then the ion is in  $\Omega_2$

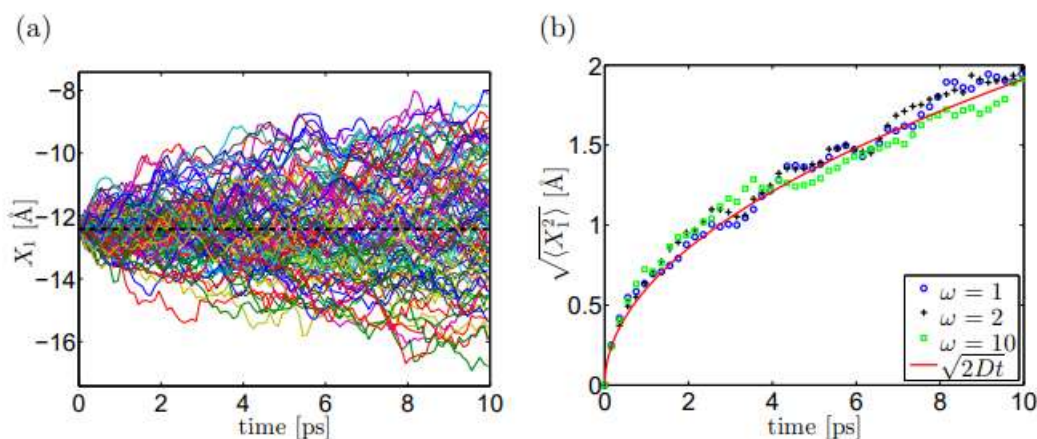


Figure 4. (a) One hundred realizations of a multiscale simulation of K<sup>+</sup> ion initiated at  $[-L/2, 0, 0]$ . We plot  $X_1$  coordinate as a function of time. Ion is described by all-atom MD for  $X_1 \geq -L/2$  and by the coarse-grained model (1.2)–(1.5) for  $X_1 < -L/2$ . The boundary between  $\Omega_1$  and  $\Omega_2$  is visualized using the black dashed line. We use  $\omega = 1$  in (6.2). (b) The mean square displacement in the first coordinate of K<sup>+</sup> ion simulated in  $\Omega_1 \cup \Omega_2$  and computed as the average of 100 realizations for  $\omega = 1$  (blue circles),  $\omega = 2$  (black crosses) and  $\omega = 10$  (green squares).

and evolves according to the coarse-grained model (1.2)–(1.5). In Figure 4(a), we use (6.2) with  $\omega = 1$  and plot  $X_1$  coordinates of all 100 realizations. We observe that the computed trajectories spread on both sides of boundary  $\partial\Omega_1$  (dashed line) without any significant bias. The mean square displacement is presented in Figure 4(b) for three different values of  $\omega$ . The results compare well with  $(2Dt)^{1/2}$  which is the mean square displacement of one coordinate of the diffusion process.

We conclude with illustrative simulations which are coupling all-atom MD with BD. We use domain  $\Omega \in \mathbb{R}^3$  decomposed into five regions as in equation (6.1), where  $\Omega_1$  and  $\Omega_2$  are given by (6.3), and

$$\Omega_3 = \left[ -\frac{L}{2} - \frac{L}{2\omega} - h_1, \frac{L}{2} + \frac{L}{2\omega} + h_1 \right]^3 \setminus (\Omega_1 \cup \Omega_2), \tag{6.4}$$

$$\Omega_4 = \left[ -\frac{L}{2} - \frac{L}{2\omega} - h_1 - h_2, \frac{L}{2} + \frac{L}{2\omega} + h_1 + h_2 \right]^3 \setminus (\Omega_1 \cup \Omega_2 \cup \Omega_3), \tag{6.5}$$

$$\Omega_5 = \mathbb{R}^3 \setminus (\Omega_1 \cup \Omega_2 \cup \Omega_3 \cup \Omega_4), \tag{6.6}$$

where  $\omega = 10$ ,  $h_1 = L/20$  and  $h_2 = L/10$ . Then the BD domain is  $\Omega_5 = \mathbb{R}^3 \setminus [-7L/10, 7L/10]^3$ . We place an ion at the origin (centre of MD domain  $\Omega_1$ ), i.e.  $X(0) = [0, 0, 0]$ , and we simulate each trajectory until it reaches the distance  $4L = 99.32 \text{ \AA}$  from the origin. Let  $T(r)$  be the time when a trajectory first reaches distance  $r$  from the origin. In Figure 5, we plot escape time  $T(r)$  as a function of distance  $r$ . We plot the value of  $T(r)$  for each realization as a blue point. The largest computed escape times (for  $r = 4L$ ) are 38, 506 ps for K<sup>+</sup> and 47, 212 ps for Na<sup>+</sup>. They are outside the range of panels in Figure 5, but the majority of data points are included in this figure. We also plot average  $\langle hT(r) \rangle$  (red solid



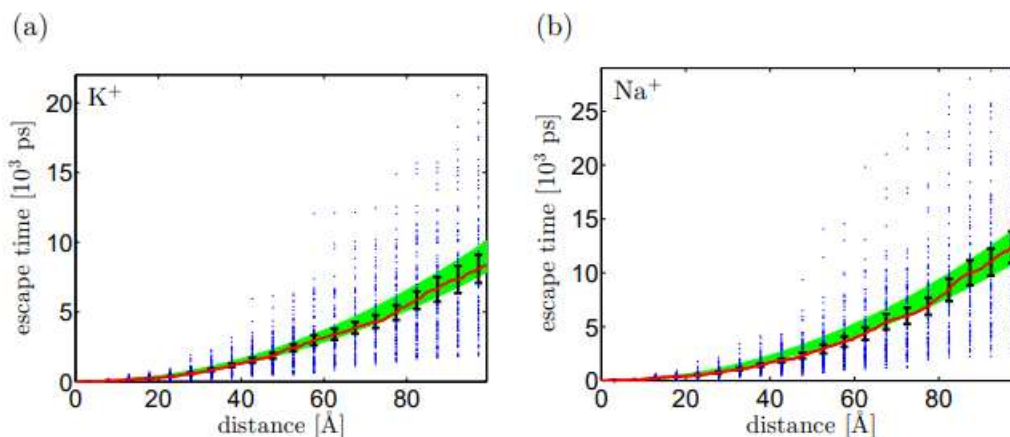


Figure 5. Escape time  $T(r)$  to reach distance  $r$  from the origin computed by the multiscale method. We consider (a)  $K^+$  ion; and (b)  $Na^+$  ion. We plot escape times for individual realizations (blue points), the mean escape time estimated from 100 realizations (red solid line) and the theoretical 95% confidence interval (6.7) (green area). We use  $\omega = 10$  in (6.2).

line) together with 95% confidence intervals. They are compared with theoretical results obtained for the BD model (1.1). The escape time distribution for the BD model (1.1) has mean equal to  $\langle T(r) \rangle = L^2 / (6D)$  and standard deviation  $L^2 / (3\sqrt{10} D)$ . The corresponding theoretical 95% confidence interval (for 100 samples) is

$$\left( \frac{L^2}{6D} - 1.96 \frac{L^2}{30D}, \frac{L^2}{6D} + 1.96 \frac{L^2}{30D} \right). \tag{6.7}$$

This interval is visualized as the green area in Figure 5. We note that it would be relatively straightforward to continue the presented multiscale computation and simulate ion diffusion in domains covering the whole cell. The most computationally intensive part is all-atom MD simulation in  $\Omega_1 \cup \Omega_2$ . However, once the ion enters  $\Omega_5$ , we can compute its trajectory very efficiently. We could further increase the BD time step in parts of  $\Omega_5$  which are far away from  $\Omega_4$ , or we could use event-based algorithms, like Green’s-function reaction dynamics [30] or First-passage kinetic Monte Carlo method [32], to compute the ion trajectory in region  $\Omega_5$

## 7. Discussion

In this paper, we have introduced and studied the coarse-grained model (1.2)– (1.5) of an ion in aqueous solution. We have parameterized this model using all-atom MD simulations for four ions ( $K^+$ ,  $Na^+$ ,  $Ca^{2+}$  and  $Cl^-$ ) and showed that this model provides an intermediate description between all-atom MD and BD simulations. It can be used both with MD time step  $\Delta t$  (to couple it with all-atom MD simulations) and BD time step  $\Delta T$  (to couple it with BD description (1.1)). In particular, the coarse-grained model enables multiscale simulations which use all-atom MD and BD in different parts of the computational domain.

In Section 6, we have illustrated this multiscale methodology using a first passage type problem where we have reported the time taken by an ion to reach a specific distance. Possible applications of this multiscale methodology include problems where a modeller

considers all-atom MD in several different parts of the cell (for example, close to binding sites or ion channels) and wants to use efficient BD simulations to transport ions by diffusion between regions where MD is used. The proposed approach thus enables the inclusion of MD-level of detail in computational domains which are much larger than would be possible to study by direct MD simulations.

y direct MD simulations. Although the illustrative simulations in Section 6 are reported over distances of the order of 102 Å, this is not a restriction of the method. Most of the computational time is spent by considering all-atom MD in  $\Omega_1 \cup \Omega_2$ . BD uses much larger time step which enables us to further extend BD region  $\Omega_5$  (and consequently, the original domain  $\Omega$ ). Moreover, if we are far away from MD domain  $\Omega_1$ , we can further increase the efficiency of BD simulations by using different BD time steps in different parts of the BD subdomain  $\Omega_5$  [12], or by using event-based BD algorithms [30, 32]. The computational intensity of BD simulations can be further decreased by using multiscale methods which efficiently and accurately combine BD models with lattice-based (compartmentbased) models [33, 34]. Such a strategy have been previously used for modelling intracellular calcium dynamics [3, 7] or actin dynamics in filopodia [35], and enables us to extend both temporal and spatial extent of the simulation.

In the literature, MD methods have been used to estimate parameters of BD simulations of ions [36]. There has also been a lot of progress in systematic coarsegraining of MD simulations [37]. The approach presented in this paper not only uses all-atom MD simulations to estimate parameters of a coarser description, but it also designs a multiscale approach where both methods are used during the same simulation. Methods which adaptively change the resolution of MD on demand have been previously reported in [38, 39]. They include algorithms which couple all-atom MD with coarse-grained MD. The coarse-grained model developed in this work does not include any water molecules and has different application areas. One of them is modelling of calcium induced calcium release through IP3R channels [3] which is discussed as a motivating example in Introduction. MD simulations in this paper use the three-site SPC/E model of water. An open question is to extend our observations and analysis to other MD models of water, which include both more detailed water models with additional sites [16, 17] and coarse-grained MD models of water [40].

## References

- [1] S. Koneshan, J. Rasaiah, M. Lynden-Bell, and S. Lee. Solvent structure, dynamics and ion mobility in aqueous solutions at 25 °C. *Journal of Physical Chemistry B*, 102:4193–4204, 1998.
- [2] M. Kohagen, P. Mason, and P. Jungwirth. Accurate description of calcium solvation in concentrated aqueous solutions. *Journal of Physical Chemistry B*, 118:7902–7909, 2014.
- [3] U. Dobramysl, S. Rüdiger, and R. Erban. Particle-based multiscale modeling of intracellular calcium dynamics. submitted, available as <http://arxiv.org/abs/1504.00146>, 2015.
- [4] B. Corry, S. Kuyucak, and S. Chung. Test of continuum theories as models of ion channels. II. Poisson-Nernst-Planck theory versus Brownian dynamics. *Biophysical Journal*, 78:2364–2381, 2000.

- [5] R. Erban, S. J. Chapman, and P. Maini. A practical guide to stochastic simulations of reaction-diffusion processes. 35 pages, available as <http://arxiv.org/abs/0704.1908>, 2007.
- [6] B. Leimkuhler and C. Matthews. *Molecular Dynamics*, volume 39 of *Interdisciplinary Applied Mathematics*. Springer, 2015.
- [7] M. Flegg, S. Rüdiger, and R. Erban. Diffusive spatio-temporal noise in a first-passage time model for intracellular calcium release. *Journal of Chemical Physics*, 138:154103, 2013.
- [8] R. Erban and S. J. Chapman. Stochastic modelling of reaction-diffusion processes: algorithms for bimolecular reactions. *Physical Biology*, 6(4):046001, 2009.
- [9] J. Lipkova, K. Zygalkis, J. Chapman, and R. Erban. Analysis of Brownian dynamics simulations of reversible bimolecular reactions. *SIAM Journal on Applied Mathematics*, 71(3):714–730, 2011.
- [10] T. Shinohara, T. Michikawa, M. Enomoto, J. Goto, M. Iwai, T. Matsu-ura, H. Yamazaki, A. Miyamoto, A. Suzuki, and K. Mikoshiba. Mechanistic basis of bell-shaped dependence of inositol 1,4,5-trisphosphate receptor gating on cytosolic calcium. *Proceedings of the National Academy of Sciences USA*, 108(37):15486–15491, 2011.
- [11] I. Serysheva. Toward a high-resolution structure of IP3R channel. *Cell Calcium*, 56:125–132, 2014. [12] R. Erban. From molecular dynamics to Brownian dynamics. *Proceedings of the Royal Society A*, 470:20140036, 2014.
- [13] H. Berendsen, J. Postma, W. Van Gunsteren, and J. Hermans. Interaction models for water in relation to protein hydration. In B. Pullman, editor, *Intermolecular Forces*, pages 331–342. D. Reidel Publishing Company, 1981.
- [14] H. Berendsen, J. Grigera, and T. Straatsma. The missing term in effective pair potentials. *Journal of Physical Chemistry*, 91(24):6169–6271, 1987.
- [15] W. Jorgensen, J. Chandrasekhar, J. Madura, R. Impey, and M. Klein. Comparison of simple potential functions for simulating liquid water. *Journal of Chemical Physics*, 79(2):926–935, 1983.