A STOCHASTIC PHASE-FIELD MODEL DETERMINED FROM MOLECULAR DYNAMICS

Erik von Schwerin[1] and Anders Szepessy[2]

Abstract. The dynamics of dendritic growth of a crystal in an undercooled melt is determined by macroscopic diffusion-convection of heat and by capillary forces acting on the nanometer scale of the solid-liquid interface width. Its modelling is useful for instance in processing techniques based on casting. The phase-field method is widely used to study evolution of such microstructural phase transformations on a continuum level; it couples the energy equation to a phenomenological Allen-Cahn/Ginzburg-Landau equation modelling the dynamics of an order parameter determining the solid and liquid phases, including also stochastic fluctuations to obtain the qualitatively correct result of dendritic side branching. This work presents a method to determine stochastic phase-field models from atomistic formulations by coarse-graining molecular dynamics. It has three steps: (1) a precise quantitative atomistic definition of the phase-field variable, based on the local potential energy; (2) derivation of its coarse-grained dynamics model, from microscopic Smoluchowski molecular dynamics (that is Brownian or over damped Langevin dynamics); and (3) numerical computation of the coarse-grained model functions. The coarse-grained model approximates Gibbs ensemble averages of the atomistic phase-field, by choosing coarse-grained drift and diffusion functions that minimize the approximation error of observables in this ensemble average.

1. Introduction to phase-field models

The phase-field model for a liquid–solid phase transformation is an Allen-Cahn/Ginzburg-Landau equation coupled to the energy equation

$$k_0\partial_t\phi = \underbrace{\text{div}(k_1\nabla\phi) - k_2\big(f'(\phi) + k_3 g'(\phi)(T - T_M)\big)}_{\partial_\phi\mathcal{F}(\phi,T)} + \text{noise}$$

$$\partial_t\big(c_v T + k_4 g(\phi)\big) = \text{div}(k_5\nabla T) \tag{1.1}$$

with a double well potential f having local minima at ± 1, smoothed step function g, temperature T, melting temperature T_M, specific heat c_v and latent heat $k_4 g(\phi)$, *cf.* [4,24]; the unknown variables to determine are

ϕ and T while the other quantities are given data. Using the phase-field variable $\phi : \mathbb{R}^3 \times [0, \infty) \to [-1, 1]$, the solid and liquid phases are interpreted as the domains $\{x \in \mathbb{R}^3 : \phi(x) > 0\}$ and $\{x \in \mathbb{R}^3 : \phi(x) < 0\}$ respectively. To have such an implicit definition of the phases, as in the level set method, is a computational advantage compared to a sharp interface model, where the necessary direct tracking of the interface introduce computational drawbacks. This phenomenological phase-field model, with free energy potentials $\mathcal{F}$ motivated by thermodynamics, has therefore become a popular and effective computational method to solve problems with complicated structures of dendrite and eutectic growth, *cf.* [1,4]. The evolution of the phase interface depends on the orientation of the solid crystal; this is modeled by an anisotropic matrix k_1. Added noise to system (1.1) is also important, *e.g.* to model nucleation phenomena, when a crystal initiates in an under-cooled liquid and starts to grow, and to obtain sidebranching dendrites [15]. The phase-field model has mathematical wellposedness [5] and convergence to sharp interface results [26]. The sharp interface limit is a consistency issue showing that, when the ratio of the diffusion term and the reaction term becomes small in the phase-field equation, the phase-field solution tends to the solution with the sharp interface of a Stefan problem. We have the situation of an almost sharp interface here since we formulate the phase-field equation in a bounded microscopic scale x, which is related to the macroscopic scale y by $x = \epsilon^{-1} y$ with a small parameter $\epsilon \ll 1$, and

$$\frac{\partial_x^2 \phi}{f(\phi)} = \epsilon^2 \frac{\partial_y^2 \phi}{f(\phi)}.$$

The aim of this work is to present a computational method to determine the data for the phase-field equation (1.1), including the noise term, from more fundamental molecular dynamics. We call the phase-field equation (1.1) a *macroscopic* model, since its variables varies on a large (macroscopic) scale; the molecular dynamics model is called a *microscopic* model here, since the particle positions variables $X : [0, \infty) \to \mathbb{R}^{3N}$ of $N \gg 1$ particles vary on a small (microscopic) scale. Phase changes can be modeled on an atomistic level by molecular dynamics or kinetic Monte Carlo methods (*i.e.* stochastic interacting particle models on a lattice as in the Ising model). To derive stochastic differential equations approximating kinetic Monte Carlo dynamics is a classical problem, studied *e.g.* in [7,20,28]. The work [16] derived coarse-grained stochastic differential equations from a kinetic Monte Carlo method with a technique related to the study here on molecular dynamics. Molecular dynamics is in some sense more attractive than kinetic Monte Carlo dynamics as a starting point for coarse-graining, since fewer parameters need to be set (none with *ab initio* molecular dynamics). Assuming that the reaction term in the Allen-Cahn equation takes a given form, for example with the common choice

$$\begin{aligned}
f(\phi) &= (1 - \phi^2)^2, \\
g(\phi) &= \frac{15}{16} \left(\frac{1}{5} \phi^5 - \frac{2}{3} \phi^3 + \phi \right) + \frac{1}{2},
\end{aligned} \tag{1.2}$$

the scalar parameters k_2, k_3, k_4, scalar functions $k_0 = k_0(\nabla\phi)$, $k_5 = k_5(\phi)$, and matrix function $k_1 = k_1(\nabla\phi)$ in the phase-field model have been *determined from atomistic molecular simulations*, using the Gibbs-Thomson condition for the temperature of a moving solid-liquid interface as a function of curvature, normal velocity, latent heat, melting temperature, the excess interface free energy, and a kinetic coefficient, *cf.* the review articles [2,12]: the excess interface free energy can be determined from molecular dynamics by introducing a "cleaving" potential to reversibly transform a solid and a liquid system to a solid-liquid system, or from position fluctuations related to interfacial stiffness; the kinetic coefficient can be determined from molecular dynamics by setting the temperature and measuring the induced velocity of an interface, or by studying the fluctuation of the number of "solid" particles. An alternative choice of phase-field functions (1.2) in [1] uses a steeper step function g to easily derive consistency with sharp interface models.

This work presents a different method to determine *data for a stochastic phase-field* equation by coarse-graining appropriate molecular dynamics – the new ingredient here is a method to determine the noise and also the actual functions in the right hand side of the phase-field equation, and not only the parameters k_i. This is made in three steps in Sections 2 to 4:

(1) a precise quantitative *atomistic definition of the phase-field* variable $m(x, X)$, based on the local potential energy of particle positions X;

(2) derivation of its coarse-grained for dynamics model $\phi(x, t) \approx m\big(X(t)\big)$, from microscopic Smoluchowski molecular dynamics (*i.e.* Brownian dynamics) of the particle positions X; and

(3) *numerical computation* of the coarse-grained model functions.

We call the approximation $\phi(x, t) \approx m\big(x, X(t)\big)$ *coarse-graining*, since the microscopic molecular description $m : \mathbb{R}^3 \times \mathbb{R}^{3N} \to \mathbb{R}$ is of much higher dimension $3N + 3 \gg 3$ than the macroscopic approximation $\phi : \mathbb{R}^3 \times [0, \infty) \to \mathbb{R}$. One may ask if the Smoluchowski (*i.e.* the over damped Langevin) dynamics is a good choice as the microscopic molecular dynamics model – our answer is that the Smoluchowski dynamics sample Gibbs ensemble averages, it generates a coarse-grained model which is a stochastic reaction-diffusion equation of the phase-field type (1.1), and Gibbs ensemble averages of the atomistic phase-field is well approximated by the coarse-grained model (in some sense optimally as explained in Sect. 4).

To determine the drift in the stochastic phase-field equation we use the setting of a traveling wave: the shape of the traveling wave, computed from molecular dynamics, determines the reaction term balancing the diffusion; since the total drift becomes zero, the reaction and diffusion terms in the phase-field equation are then determined up to a multiplicative constant – this multiplicative constant is chosen to match equilibrium fluctuations of the phase-field. The diffusion coefficient for the noise in the stochastic phase-field is determined from an optimization problem, where the diffusion is chosen to minimize the error of observables in the Gibbs ensemble.

This alternative method to determine the phase-field model from atomistic simulations comes with some new ideas, *e.g.* handling the noise, but it would need further verification and comparison by others to reach similar maturity as the established methods in [2] for atomistic verification of phase-field parameters.

We present computational results for an example based on Lennard-Jones type molecular dynamics in an equilibrium simulation with constant number of particles, volume, and temperature and with periodic boundary conditions; we compute both the noise function and the free energy function at the melting temperature, *i.e.* $\mathcal{F}(\phi, T_M)$, for some orientations of the solid-liquid interface with respect to the crystal. To determine the free energy function above and below the melting temperature is also important; that requires a different molecular dynamics simulation in a non periodic setting following a moving interface; see Remark 5.1. Other relevant extensions for the future would be to include constant pressure and hybrid simulations, see Remarks 5.1 and 5.2.

2. QUANTITATIVE ATOMISTIC DEFINITION OF THE PHASE-FIELD VARIABLE

The aim is to give a unique definition of the phase-field variable, so that it can be determined precisely from atomistic simulations. The usual interpretation is to measure interatomic distances or to use radial distribution functions (*cf.* Fig. 4) to measure where the phase is solid and where it is liquid, which then implicitly defines the phase-field variable [4]. Here we instead use the energy equation for a quantitative and explicit definition of the phase-field variable. The macroscopic energy equation with a phase transformation and heat conduction is

$$\partial_t(c_v T + m) = \text{div}(k \nabla T) \tag{2.1}$$

where m corresponds to the latent heat release. In (1.1) the latent heat determines the parameter k_2, since ϕ is defined to jump from 1 to -1 in the phase transformation. We will instead use this latent heat to directly define the phase-field function, and not only the parameter k_2. The latent heat is the change in enthalpy of a phase transformation at constant pressure, see [18,19,22]. We will study molecular dynamics at constant volume,
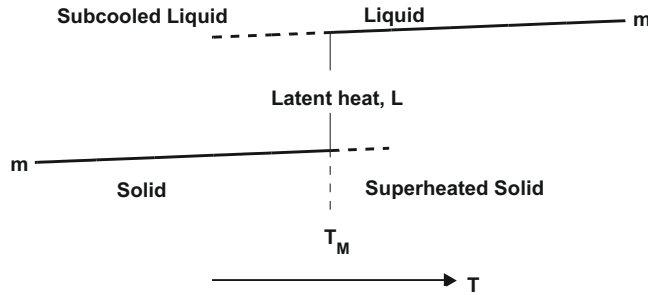
FIGURE 1. Schematic picture of $m(T)$ for a pure liquid (top curve) and a pure solid (bottom curve) and the latent heat as the jump in m at a phase transition.

instead of pressure; then the latent heat is the change in the potential energy. The total energy, $c_v T + m$, can be defined from molecular dynamics of N particles with position X_i, velocity v_i, and mass μ in a potential V,

$$c_v T + m = \sum_{i=1}^{N} \mu \frac{|v_i|^2}{2} + V(X_1, \ldots, X_N), \tag{2.2}$$

see [11,13]. We assume for simplicity that the potential is defined from pair interactions

$$V(X) = \frac{1}{2} \sum_{i=1}^{N} \sum_{j \neq i} \Phi(X_i - X_j), \tag{2.3}$$

where $\Phi : \mathbb{R}^3 \to \mathbb{R}$ is a molecular dynamics pair potential, $e.g.$ a Lennard-Jones potential

$$\Phi(x) = z_1 \left(\frac{\sigma}{|x|} \right)^{12} - z_2 \left(\frac{\sigma}{|x|} \right)^{6};$$

in this work we use the exponential-6 pair potential (modelling Argon at high pressure), which is almost the same as the Lennard-Jones potential, see [29]. In the macroscopic setting the jump of m in a phase change is called the latent heat, which depends on the thermodynamic variables kept constant: with constant N, T, and volume it is called the internal energy and with constant pressure instead of volume it is called enthalpy. Here we study the ensemble with constant volume; Remark 5.1 comments on the constant pressure ensemble, which is more common in applications. For a molecular system in equilibrium with constant volume, the mean of the kinetic energy, $\sum_i \mu |v_i|^2/2$, is N times the temperature. It is therefore natural to let the phase-field variable be determined by the potential energy $V(X)$, which typically has a jump at the melting temperature that we identify with the latent heat; see Figure 1. In a point-wise setting the potential energy can be represented by the distribution

$$\frac{1}{2} \sum_{i=1}^{N} \sum_{j \neq i} \Phi(X_i - X_j) \delta(x - X_i)$$

where δ is the point mass at the origin [13]. We seek an averaged variant and we will study a microscopic phase change model where the interface is almost planar in the microscopic scale with normal in the x_1 direction.

Therefore we take a smooth average and define the phase-field variable by

$$m(X, x) := \frac{1}{2} \sum_{i=1}^{N} \underbrace{\sum_{j \neq i} \Phi(X_i - X_j)\, \eta(x - X_i)}_{m_i(X)} \tag{2.4}$$

where $\eta : \mathbb{R}^3 \to (0, \infty)$ is a smooth approximation of the point (delta) mass, with scale $\epsilon_i > 0$ in the x_i direction,

$$\eta(x) := \prod_{i=1}^{3} \frac{\mathrm{e}^{-|x_i|^2/(2\epsilon_i^2)}}{(2\pi\epsilon_i^2)^{1/2}}. \tag{2.5}$$

Smooth averages have been used in molecular dynamics for fluid dynamics, *cf.* [11], and for the vortex blob method and the smoothed particle hydrodynamics approximation of moving particles in fluid dynamics, *cf.* [3,21]. Sections 3 and 4 present a molecular dynamics model for the potential energy (2.4).

We assume our molecular system is in local equilibrium, meaning that we want to determine Gibbs ensemble averages of the defined phase-field variable

$$\frac{\int_{\mathbb{R}^{3N}} g\big(m(X, x)\big) \mathrm{e}^{-V(X)/(k_{\mathrm{B}} T)}\, \mathrm{d}X}{\int_{\mathbb{R}^{3N}} \mathrm{e}^{-V(X)/(k_{\mathrm{B}} T)}\, \mathrm{d}X} \tag{2.6}$$

for any function $g : \mathbb{R} \to \mathbb{R}$, where the temperature, T, may depend on x in the macroscopic scale and k_{B} is the Boltzmann constant. In the next section we determine a coarse-grained model that approximates such Gibbs averages.

3. An atomistic Smoluchowski dynamics model

The standard method to simulate molecular dynamics in the microcanonical ensemble of constant volume, energy, and number of particles is to write Newton's laws for the particles, *cf.* [9,25]. Since we are only interested in the Gibbs equilibrium averages (2.6), not depending explicitly on time, an alternative is instead to use Smoluchowski dynamics at constant temperature, T, with the Ito differential equations

$$\mathrm{d}X_i^t = -\partial_{X_i} V(X^t)\mathrm{d}t + \sqrt{2\gamma}\, \mathrm{d}W_i^t \tag{3.1}$$

where W_i are independent Brownian motions, $X_i^t := X_i(t)$ is the position of the ith particle at time t and $\gamma := k_{\mathrm{B}} T$. Here, the important property of the Smoluchowski dynamics is that its invariant measure is the Gibbs measure

$$\frac{\mathrm{e}^{-V(X)/\gamma}\, \mathrm{d}X}{\int_{\mathbb{R}^{3N}} \mathrm{e}^{-V(X)/\gamma}\, \mathrm{d}X} =: G(X)\mathrm{d}X, \tag{3.2}$$

under appropriate assumptions on V, *cf.* [6].

The Smoluchowski equation is the zero relaxation time limit (*i.e.* $\tau \to 0+$) of Langevin's equation (*cf.* [14, 17,23,25])

$$\mathrm{d}X_i^s = v_i^s\, \mathrm{d}s,$$

$$M\mathrm{d}v_i^s = -\partial_{X_i} V(X^s)\, \mathrm{d}s - \frac{v_i^s}{\tau}\, \mathrm{d}s + \sqrt{\frac{2\gamma}{\tau}}\, \mathrm{d}\hat{W}_i^s, \tag{3.3}$$

in the faster time scale $s = t/\tau$, where M is the mass and $\hat{W}_i$ are independent Brownian motions, and the standard Hamiltonian molecular dynamics corresponds to vanishing dissipation and fluctuation (by taking $\tau = \infty$) in the Langevin dynamics.
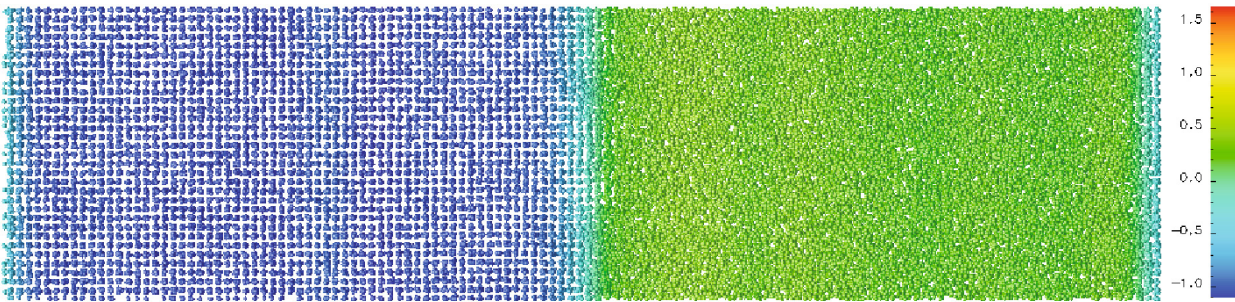
FIGURE 2. A configuration of a Smoluchowski dynamics two-phase system. The left part is solid and the right part liquid. The computational cell has periodic boundary conditions in all three directions so that the cell represents part of a layered structure of alternating infinite slabs of solid and liquid. Here the solid-liquid interface is in the {100} plane of the face centered cubic crystal in the solid. The color of the particles represents the phase-field computed as discrete time averages approximating $\frac{1}{t_2-t_1}\int_{t_1}^{t_2} m(x_1; X^t)\,\mathrm{d}t$; see Figure 3(a).

Our *microscopic model* of a phase change is then the Smoluchowski dynamics model (3.1) with the phase-field variable m in (2.4) coupled to the macroscopic energy equation (2.1). The temperature varies on the macroscopic scale, due to the energy equation, so that T is almost constant on the microscopic scale of a molecular dynamics simulation and makes its Gibbs equilibrium density proportional to $\mathrm{e}^{-V(X)/(k_\mathrm{B}T(x))}$ reasonable. The reason we use Smoluchowski dynamics is that:

- it samples the Gibbs ensemble averages (2.6) time asymptotically; and
- it is a dynamics that generates a coarse-grained equation of the same qualitative type as the phase-field equation (1.1), namely a stochastic reaction-diffusion partial differential equation.

Figures 2–4 show the phase-field, the density and the radial distribution functions for a phase transition in equilibrium, at the melting temperature using the exponential-6 pair potential [29,30] and the Smoluchowski dynamics (3.1). Since the Smoluchowski dynamics samples the Gibbs ensemble, these figures would be the same for any other molecular dynamics method sampling this Gibbs ensemble, as *e.g.* the Langevin dynamics (3.3) with any positive relaxation time τ.

4. COARSE-GRAINED PHASE-FIELD DYNAMICS

We want to determine a mean drift function $a(\bar{m})$ and a diffusion function $b(\bar{m})$ so that the coarse-grained approximation $\bar{m}^t$, solving the coarse-grained equation

$$\mathrm{d}\bar{m}^t = a(\bar{m}^t)\mathrm{d}t + \sum_{k=1}^{M} b_k(\bar{m}^t)\mathrm{d}\tilde{W}_k^t, \tag{4.1}$$

optimally approximates the Gibbs ensemble averages (2.6), of the phase-field $m(X^t, \cdot)$ defined in (2.4) with X^t solving the Smoluchowski dynamics (3.1). The Brownian motions $\tilde{W}_k$, $k = 1, \ldots, M$, are mutually independent and also independent of all W_i. To obtain an optimal approximation, we seek the minimal error of the Gibbs ensemble average

$$\min_{a,b} \lim_{\mathcal{T}\to\infty} \mathcal{T}^{-1} \int_0^{\mathcal{T}} \mathbb{E}\big[g\big(m(X^\mathcal{T}, \cdot)\big)\big] - \mathbb{E}\big[g(\bar{m}^\mathcal{T})\big]\mathrm{d}\mathcal{T} \tag{4.2}$$
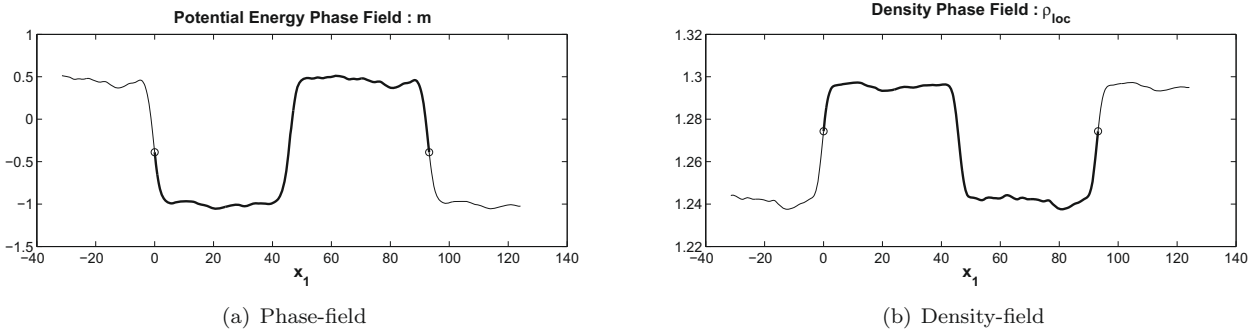
(a) Phase-field

(b) Density-field

FIGURE 3. Figure (a) shows a phase-field computed as a discrete time average approximating $\frac{1}{t_2-t_1}\int_{t_1}^{t_2} m(x_1; X^t)\,\mathrm{d}t$, where X^t follows the Smoluchowski dynamics. The spatial averages are taken over the entire width of the computational cell in planes parallel to the interfaces, while remaining localised, with $\epsilon_1 = 1$, in the direction normal to the interfaces. The thick curve shows the phase-field over one period in the layered structure; as a consequence of the periodic boundary conditions the computational cell contains two interfaces. Figure (b) shows the corresponding computed density using the particle density instead of the potential energy; see Remark 4.2.

for any given function $g : \mathbb{R} \to \mathbb{R}$, with the same initial value $\bar{m}^0 = m(X^0, \cdot)$. Such minimizing functions, a and b, lead (by ergodicity) to our objective – to approximate the Gibbs average defined in (2.6)

$$\lim_{\mathcal{T}\to\infty} \mathcal{T}^{-1} \int_0^{\mathcal{T}} \mathbb{E}\big[g\big(m(X^{\mathcal{T}}, \cdot)\big)\big]\mathrm{d}\mathcal{T} = \frac{\int_{\mathbb{R}^{3N}} g\big(m(X, x)\big)\mathrm{e}^{-V(X)/(k_{\mathrm{B}}T)}\,\mathrm{d}X}{\int_{\mathbb{R}^{3N}} \mathrm{e}^{-V(X)/(k_{\mathrm{B}}T)}\,\mathrm{d}X}.$$

We perform this optimal control way of coarse-graining in three steps, described below. Note that there may be other microscopic dynamics with different coarse-graining functions that lead to smaller approximation error – our optimality concerns coarse-grained drift and diffusion functions a and b for the given microscopic Smoluchowski dynamics (3.1).

The first idea in the coarse-graining procedure, in Section 4.1, is that *Ito's formula and the Smoluchowski dynamics* (3.1) determine functions α and β, depending on the microscopic state X, so that

$$\mathrm{d}m(X^t, \cdot) = \alpha(X^t)\mathrm{d}t + \sum_{j=1}^{N} \beta_j(X^t)\mathrm{d}W_j^t. \tag{4.3}$$

The next step, in Section 4.2, is to *estimate the error*, using the Kolmogorov equations for $\bar{m}$ and (4.3), similar to [16,27], which leads to

$$\mathbb{E}\big[g\big(m(X^{\mathcal{T}}, \cdot)\big) - g(\bar{m}^{\mathcal{T}})\big] = \mathbb{E}\left[\int_0^{\mathcal{T}} \langle \bar{u}', \alpha - a\rangle + \left\langle \bar{u}'', \sum_{j=1}^{N} \beta_j \otimes \beta_j - \sum_{k=1}^{M} b_k \otimes b_k \right\rangle \mathrm{d}t\right],$$

where $\langle \bar{u}', \cdot \rangle$ is the $L^2(\mathbb{R}^3)$ scalar product (corresponding to the variable $x \in \mathbb{R}^3$) with $\bar{u}'$, which is the Gateaux derivative (*i.e.* functional derivative) of the functional $\mathbb{E}[g(\bar{m}^{\mathcal{T}})\,|\,\bar{m}^t = n]$ with respect to n; and similarly $\langle \bar{u}'', \cdot \rangle$ is the $L^2(\mathbb{R}^3 \times \mathbb{R}^3)$ scalar product with the second Gateaux derivative $\bar{u}''$ of the functional $\mathbb{E}[g(\bar{m}^{\mathcal{T}})\,|\,\bar{m}^t = n]$ with respect to n. The notation $b_k \otimes b_k(x, x') := b_k(x)b_k(x')$ is the tensor product.
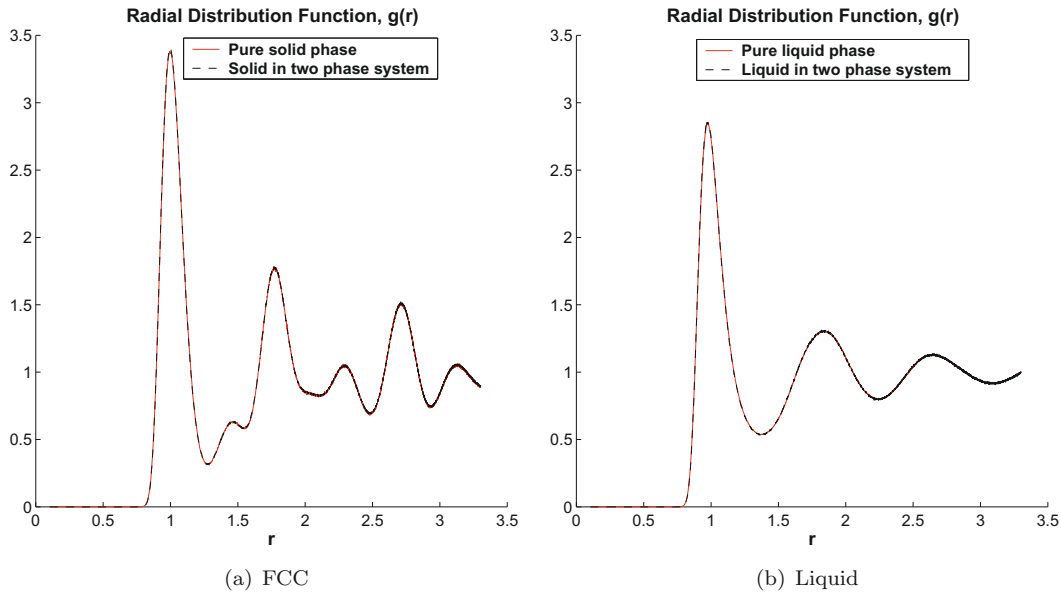
FIGURE 4. The radial distribution function, $g(r)$, computed from several configurations, separated in time, in the process of setting up one of the two-phase systems. The solid curve shows $g(r)$ computed as an average over all particles in the computational cell used while presimulating the solid and the liquid part, in (a) and (b) respectively. The dashed curves show $g(r)$ computed as an average over particles in two slices of the computational cell of the two-phase system; (a) shows $g(r)$ obtained from a slice inside the solid phase, and (b) shows $g(r)$ from a slice inside the liquid phase. The radial distribution functions show good agreement between the single phase systems and the corresponding solid and liquid subdomains away from the interface.

The final step, in Section 4.3, is to use molecular dynamics simulations for a planar interface two phase problem and *compute averages* in cross sections parallel to the interface, where $\bar{u}'$, $\bar{u}''$, a, and $\sum_k b_k \otimes b_k$ are $\bar{m}$-independent, to evaluate approximations to the functions a and $\sum_k b_k \otimes b_k$ by

$$a = \lim_{\mathcal{T} \to \infty} \frac{1}{\mathcal{T}} \mathbb{E}\left[\int_0^{\mathcal{T}} \alpha \mathrm{d}t \right],$$

$$\sum_k b_k \otimes b_k = \lim_{\mathcal{T} \to \infty} \frac{1}{\mathcal{T}} \mathbb{E}\left[\int_0^{\mathcal{T}} \sum_{j=1}^N \beta_j \otimes \beta_j \mathrm{d}t \right].$$

$$(4.4)$$

Ergodicity implies that the approximations can be written as the Gibbs ensemble averages

$$a(x) = \int_{\mathbb{R}^{3N}} \alpha(x; X) G(X) \mathrm{d}X$$

$$\sum_k b_k \otimes b_k(x, y) = \int_{\mathbb{R}^{3N}} \beta_j \otimes \beta_j(x, y; X) G(X) \mathrm{d}X,$$

$$(4.5)$$

with the Gibbs density G defined in (3.2).

4.1. **The Ito formula for the phase-field**

The Ito formula (*cf.* [10]) implies

$$\mathrm{d}m(X^t, x) = \underbrace{\sum_{j=1}^{N}(-\partial_{X_j}m\,\partial_{X_j}V + \gamma\partial_{X_jX_j}m)\,\mathrm{d}t}_{\alpha(X^t)} + \sum_{j=1}^{N}\underbrace{\sqrt{2\gamma}\partial_{X_j}m\,\mathrm{d}W_j}_{\beta_j(X^t)}. \tag{4.6}$$

The definition in (2.4),

$$m(X^t, x) = \sum_i m_i(X)\eta(x - X_i^t),$$

yields

$$\partial_{X_j}m = \sum_i \partial_{X_j}m_i\eta(x - X_i) + m_j\partial_{X_j}\eta(x - X_j).$$

In (4.6) we will use (2.5) to evaluate the last derivative as

$$\partial_{X_j}\eta(x - X_j) = -\partial_x\eta(x - X_j), \qquad\qquad \text{in d}t \text{ terms,}$$

$$\partial_{X_j}\eta(x - X_j) = -\eta(x - X_j)\left(\frac{(x - X_j)_1}{\epsilon_1^2}, \frac{(x - X_j)_2}{\epsilon_2^2}, \frac{(x - X_j)_3}{\epsilon_3^2}\right), \qquad \text{in d}W_j \text{ terms,}$$

in order to avoid spatial derivatives on the diffusion coefficient, while including them in the drift. Since

$$m_i = \frac{1}{2}\sum_{k \neq i}\Phi(X_i - X_k)$$

and

$$V(X) = \frac{1}{2}\sum_i\sum_{j \neq i}\Phi(X_i - X_j)$$

there holds

$$\partial_{X_j}m_i = \frac{1}{2}\sum_{k \neq i}\Phi'(X_i - X_k)\delta_{ij} - \frac{1}{2}\Phi'(X_i - X_j)(1 - \delta_{ij}),$$

$$\partial_{X_j}V(X) = \sum_{k \neq j}\Phi'(X_j - X_k),$$

where

$$\delta_{ij} := \begin{cases} 1 & i = j, \\ 0 & i \neq j, \end{cases}$$

is the Kronecker symbol. The second derivatives are

$$\partial_{X_jX_j}m = \sum_i \partial_{X_jX_j}m_i\eta(x - X_i) - 2\partial_{X_j}m_j\partial_x\eta(x - X_j) + m_j\partial_{xx}\eta(x - X_j),$$

with

$$\partial_{X_jX_j}m_i = \frac{1}{2}\sum_{k \neq i}\Phi''(X_i - X_k)\delta_{ij} + \frac{1}{2}\Phi''(X_i - X_j)(1 - \delta_{ij}).$$

and all terms in (4.6) are now expressed in terms of Φ, its gradient Φ' and Hessian Φ''. We note that the drift, α, has the form

$$
\alpha = \partial_{xx}A_2 + \partial_x A_1 + A_0 :=
$$
$$
= \gamma\partial_{xx}m(X^t,x) + \partial_x\left(\sum_{i=1}^{N} n_{1i}(X^t)\eta(x-X_i^t)\right) + \sum_{i=1}^{N} n_{0i}(X^t)\eta(x-X_i^t) \tag{4.7}
$$

of conservative and non conservative reaction terms. Similarly the diffusion, β_j, takes the form

$$
\sum_{i=1}^{N} n_{3ji}(X^t)\eta(x-X_i^t) + n_{4j}(X^t)\eta(x-X_j^t)(x-X_j^t).
$$

4.2. The error representation

The conditioned expected value
$$
\bar{u}(n,t) := \mathbb{E}[g(\bar{m}^{\mathcal{T}})\,|\,\bar{m}^t = n] \tag{4.8}
$$
satisfies the Kolmogorov equation (*cf.* [16,27])

$$
\partial_t u + \langle \bar{u}', a\rangle + \left\langle \bar{u}'', \sum_{k=1}^{M} b_k \otimes b_k \right\rangle = 0,
$$
$$
\bar{u}(\cdot, \mathcal{T}) = g. \tag{4.9}
$$

Let $m^t := m(X^t,\cdot)$. The final condition in (4.9) and the definition (4.8) show that

$$
\mathbb{E}\Big[g\big(m(X^{\mathcal{T}},\cdot)\big) - g(\bar{m}^{\mathcal{T}})\Big] = \mathbb{E}[\bar{u}(m^{\mathcal{T}},\mathcal{T})] - \bar{u}(m^0,0) = \mathbb{E}\left[\int_0^{\mathcal{T}} \mathrm{d}\bar{u}(m^t,t)\right].
$$

Use the Ito formula and (4.6) to evaluate $\mathrm{d}\bar{u}(m^t,t)$ and then Kolmogorov's equation (4.9) to replace $\partial_t\bar{u}$ in this right hand side to obtain the error representation

$$
\mathbb{E}\big[g\big(m(X^{\mathcal{T}},\cdot)\big) - g(\bar{m}^{\mathcal{T}})\big] = \mathbb{E}\left[\int_0^{\mathcal{T}} \langle \bar{u}', \alpha\rangle + \left\langle \bar{u}'', \sum_{j=1}^{N} \beta_j \otimes \beta_j \right\rangle + \partial_t\bar{u}\ \mathrm{d}t\right]
$$
$$
= \mathbb{E}\left[\int_0^{\mathcal{T}} \langle \bar{u}', \alpha - a\rangle + \left\langle \bar{u}'', \sum_{j=1}^{N} \beta_j \otimes \beta_j - \sum_{k=1}^{M} b_k \otimes b_k \right\rangle\ \mathrm{d}t\right].
$$

Remark 4.1 (energy fluctuation)**.** If we integrate the noise term over all x_1, *i.e.* take ϵ_1 very large, and let $g(m) = m^2$, then the error we are studying $\mathbb{E}\big[g\big(m(X^{\mathcal{T}},\cdot)\big) - g(\bar{m}^{\mathcal{T}})\big]$ is the usual fluctuation of energy $\mathbb{E}[V^2 - \mathbb{E}[V]^2]$ (proportional to the specific heat [14]), provided we set $\bar{m} = \mathbb{E}[V]$.

4.3. Computation of averages in cross sections

The optimal choice of the function $b \otimes b$ is to minimize $\mathbb{E}[\int_0^{\mathcal{T}} \langle \bar{u}'', \beta \otimes \beta - b \otimes b\rangle \mathrm{d}t]$, which seems hard to determine exactly, since $\bar{u}''\big(m(X^t,\cdot),t\big)$ depends on X^t. However, the function $\bar{u}''\big(m(X^t,\cdot),t\big)$ depends only on the coarse-grained $m(X^t,\cdot)$ and not directly on X^t and not on t for $\mathcal{T} \to \infty$. A possible approximation is to neglect the fluctuations of $m(X^t,\cdot)$ and set $m(X^t,x) = \bar{m}(x)$ in the ϵ-neighbourhood of x and write

$$
\mathbb{E}\left[\int_0^{\mathcal{T}} \langle \bar{u}''\big(m(X^t,x),t\big), \beta \otimes \beta(x) - b \otimes b\big(\bar{m}(x)\big)\rangle \mathrm{d}t\right] \approx \left\langle \bar{u}''\big(\bar{m}(x),\mathcal{T}\big), \mathbb{E}\left[\int_0^{\mathcal{T}} \beta \otimes \beta(x) - b \otimes b\big(\bar{m}(x)\big)\mathrm{d}t\right]\right\rangle
$$

which leads to the minimizing condition

$$\lim_{\mathcal{T}\to\infty} \mathcal{T}^{-1}\mathbb{E}\left[\int_0^{\mathcal{T}} \beta\otimes\beta(x) - b\otimes b\big(\bar{m}(x)\big)\mathrm{d}t\right] = 0,$$

and means that the diffusion matrix $\overline{d}(x,x') := b\otimes b\big(\bar{m}(\cdot,x)\big)$ is

$$\overline{d}(x,x') = \lim_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}}\mathbb{E}\left[\int_0^{\mathcal{T}} \sum_{j=1}^{N} \beta_j(x)\otimes\beta_j(x')\mathrm{d}t\right], \tag{4.10}$$

and similarly the drift $\overline{a}(x) := a(\bar{m}(\cdot,x))$ is

$$\overline{a}(x) = \lim_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}}\mathbb{E}\left[\int_0^{\mathcal{T}} \alpha(x)\mathrm{d}t\right]. \tag{4.11}$$

In other words, it is useful to think of an expansion of $\bar{u}'$ in $\alpha - a$ and determine a by the leading order conditions (4.11) and (4.10).

We expect the spatial averages of the microscopic variables to vary on a much smaller scale in the x_1 direction normal to the phase front than in its orthogonal directions; consequently we use an average function η in (2.4) with higher resolution in the x_1 direction, so that $0 < \epsilon_1 \ll \epsilon_2 = \epsilon_3$. In a microscopic simulation the molecular dynamics (3.1) has a small spatial volume, so that ϵ_2 is much larger than the size of the simulation box. Consequently we may first think of α and β depending only on the x_1 coordinate.

In practice, diffusion $\overline{d}$ can only be determined for a discrete set of points

$$\{(x_1(1), x_2(1), x_3(1)), \ldots, (x_1(M), x_2(M), x_3(M))\} =: \mathcal{X}_M$$

and $\mathcal{X}_M \times \mathcal{X}_M$, respectively, related to the scales ϵ_i. The diffusion coefficient $\overline{b}$, as a function of x, is then obtained as the square root of the $M\times M$ matrix $\overline{d}$. We expect that $x_1 \mapsto \mathcal{T}^{-1}\mathbb{E}[\int_0^{\mathcal{T}} m^t\mathrm{d}t]$ is monotone in the interface, for fixed (x_2,x_3), so that its inverse function, denoted by m^{-1}, is well defined. Then the coarse-grained drift and the diffusion can be obtained as functions of $\bar{m}$ by

$$d(\bar{m}) := \overline{d}\big(m^{-1}(\bar{m})\big),$$

and similarly for a. The Gibbs ensemble average and (4.6), (3.2) show that the drift vanishes

$$
\begin{aligned}
\overline{a} &= \int_{\mathbb{R}^{3N}} \alpha(X)G(X)\mathrm{d}X \\
&= \int_{\mathbb{R}^{3N}} \sum_{j=1}^{N} (-\partial_{X_j}m\,\partial_{X_j}V + \gamma\partial_{X_jX_j}m)\,G(X)\mathrm{d}X \\
&= \int_{\mathbb{R}^{3N}} \sum_{j=1}^{N} \big(\gamma\partial_{X_j}m\,\partial_{X_j}G(X) + \gamma\partial_{X_jX_j}m\,G(X)\big)\mathrm{d}X \\
&= \int_{\mathbb{R}^{3N}} \sum_{j=1}^{N} \big(-\gamma\partial_{X_jX_j}m\,G(X) + \gamma\partial_{X_jX_j}m\,G(X)\big)\mathrm{d}X \\
&= 0.
\end{aligned}
$$

That is, in an equilibrium system the total drift vanishes and consists of two balancing terms. We will use the setting of a traveling wave of m to determine the drift, which we define as the diffusion-reaction operator

$$a = \sigma\big(T\partial_{xx}\bar{m} - A_0(\bar{m})\big).$$

The (somewhat ad hoc) motivation for this diffusion-reaction form is that the phase-field model (1.1) has such form and that the atomistic phase-field yields a drift $\alpha(X)$ with a diffusion and a reaction term (with two and zero spatial derivatives, respectively). We avoid a possible convection term with one spacial derivative, since a macroscopic velocity field is inappropriate for the model (1.1). We write the diffusion term simply as $\sigma\partial_{xx}\bar{m}$ and do not attempt to determine the functions in a general diffusion term $\sigma_1\partial_x(\sigma_2\partial_x\bar{m})$. We find the reaction term A_0 and the mobility factor σ in two steps: first the traveling wave $m(x) = m(X, x)$, computed from molecular dynamics, is used to approximate $\bar{m}(x)$ and the reaction term is determined by imposing

$$\overline{A}_0(x) = T\partial_{xx}m(x),$$

and letting $A_0(\bar{m}) := \overline{A}_0\big(m^{-1}(\bar{m})\big)$. Then the positive function $\sigma(\bar{m})$ is chosen to match the variance of the potential energy (as liquid and as solid phase) – we have

$$\mathrm{d}\bar{m} = \sigma\big(T\partial_{xx}\bar{m} - A_0(\bar{m})\big)\mathrm{d}t + \sum_k b\ \mathrm{d}\tilde{W}_k \tag{4.12}$$

and define the spacial average

$$M := \int_U \bar{m}\mathrm{d}x$$

to obtain (from integration, linearization, and periodic boundary conditions) the Ornstein-Uhlenbeck process

$$\mathrm{d}M = -\sigma(\bar{m})A_0'(\bar{m})M\mathrm{d}t + \int_U \sum_k b(\bar{m})\ \mathrm{d}\tilde{W}_k\mathrm{d}x.$$

The function $\sigma(\bar{m})$ can be determined (for the liquid and solid state) by equating the variance of M and the variance of the spatial average of the atomistic phase-field $\int_U m(x; X)\mathrm{d}x$. Figure 5(b) shows that the potential k_2f, defined by $k_2f'(\bar{m}) = A_0(\bar{m})$, is convex in a neighborhood of the equilibrium points (where $A_0'(\bar{m}) = 0$), so that $A_0'(\bar{m}) > 0$ and the invariant distribution of M is a normal distribution with mean zero and variance proportional to $1/(\sigma A_0')$.

Remark 4.2 (density dynamics)**.** A similar derivation with Ito's formula shows that the density, $\rho(x) := \sum_i \eta(x - X_i)$, satisfies

$$\mathrm{d}\rho = \gamma\partial_{xx}\rho\ \mathrm{d}t - \partial_x \sum_{i=1}^N \eta(x - X_i)\big(-V'(X)\ \mathrm{d}t + \sqrt{2\gamma}\ \mathrm{d}W_i\big).$$

The expression for the coarse-grained values (4.4) implies that the drift for the density satisfies

$$\gamma\partial_{xx}\underbrace{\int_{\mathbb{R}^{3N}} \sum_i \eta(x - X_i)G(X)\mathrm{d}X}_{=:\bar{\rho}} - \partial_x\int_{\mathbb{R}^{3N}} \sum_{i=1}^N \eta(x - X_i)\underbrace{\big(-V'(X)\big)G(X)}_{\gamma\partial_X G(X)}\mathrm{d}X$$

$$= \gamma\partial_{xx}\bar{\rho} - \gamma\partial_{xx}\int_{\mathbb{R}^{3N}} \sum_{i=1}^N \eta(x - X_i)G(X)\mathrm{d}X = 0$$

which leads to a stochastic reaction diffusion equation of the same qualitative form as for the phase-field $\bar{m}$.

(a) $-\overline{A_0}(x) = \gamma\partial_x^2\bar{m}(x)$

(b) $k_2 f(m) = \int_{m_0}^m -\overline{A_0}(x(\tilde{m}))\,d\tilde{m}$
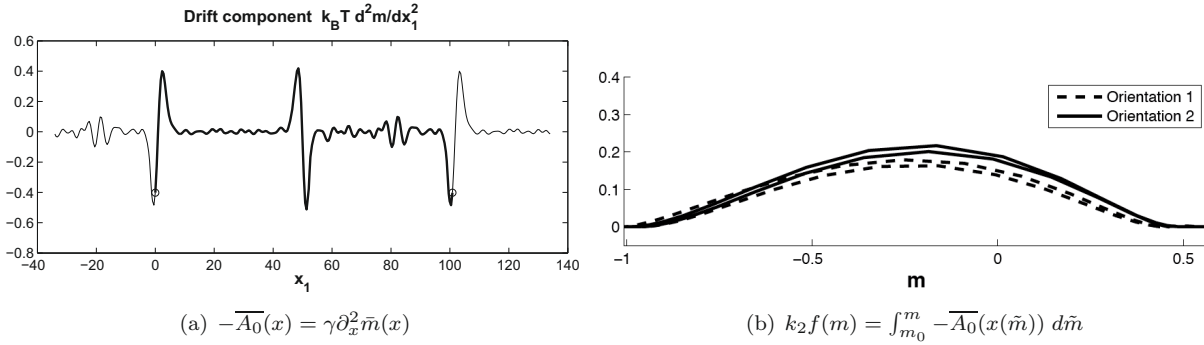
FIGURE 5. Figure (a) shows an example of a computed reaction term $-\overline{A_0}(x)$ over the periodic computational cell containing two interfaces. Inverting the computed phase-field function inside the interfaces and integrating with respect to m gives the double-well type potentials shown in (b) for two different orientations of the solid-liquid interface with respect to the face centered cubic crystal structure in the solid; in Orientation 1, O_1, the interface is along the $\{100\}$ plane in the crystal and in Orientation 2, O_2, it is along the $\{111\}$ plane. The two curves with the same orientation correspond to the two phase transitions in the computational domain, explained in Figure 3. Note that the potentials are indeed qualitative similar to the function $\phi \mapsto k_2(1-\phi^2)^2$ in (1.2) if we rescale to have the equilibrium points in $m = \pm 1$ and we restrict the domain to $\phi \in [-1, 1]$ where the phase-field has it values.

4.4. Computed coarse-grained phase-field

Figure 5(a) shows the computed reaction term $\overline{A_0}$ for one particular orientation of the solid crystal with respect to the phase interface. By inverting the computed function $x_1 \mapsto \mathcal{T}^{-1}\mathbb{E}[\int_0^{\mathcal{T}} m^t dt]$ locally in the domain of monotonicity at the interface and integrating $\overline{A_0}(x(m))$ with respect to m we obtain the expected double well shape of the potential $k_2 f(m)$; Figure 5(b) shows the computed reaction function $k_2 f(m)$ for two different orientations of the solid crystal with respect to the phase interface, *i.e.* both the coefficient k_2 and the function f in (1.2) are determined together. We want to write the coarse-grained equation (4.12) for $\bar{m}$ in the same form as (1.1). That allows orientation dependent factors in the diffusion term, the time derivative and the noise term but not in the reaction term. Therefore we will now factor out the orientation dependence in the reaction term in (4.12). Let $q_i := \overline{A_0(O_i)}/\overline{A_0(O_1)}$ be the quotient of the two reaction terms, which is the orientation dependent factor in the reaction term (and by (4.13) determines the orientation dependence in k_1). Also the time scale may depend on the crystal orientation since the relaxation time τ may differ for the two orientations; we have $s = t/\tau$, so let $r_i := \tau(O_i)/\tau(O_1)$. Rescale the time in the computation with orientation O_i, to be the same as for orientation O_1, to obtain

$$\mathrm{d}\bar{m}_i = r_i q_i\big(\gamma q_i^{-1}\partial_x^2\bar{m}_i + \overline{A_0}(\bar{m}_i; O_1)\big)\mathrm{d}t + r_i^{1/2}\sum_k \bar{b}_k(\bar{m}_i; O_i)\mathrm{d}\tilde{W}_k^t$$

which can be written in the same form as (1.1)

$$(r_i q_i)^{-1}\mathrm{d}\bar{m} = \big(\gamma q_i^{-1}\partial_x^2\bar{m} + \overline{A_0}(\bar{m}; O_1)\big)\mathrm{d}t + r_i^{-1/2}q_i^{-1}\sum_k \bar{b}(\bar{m}; O_i)\mathrm{d}\tilde{W}_k^t \qquad (4.13)$$
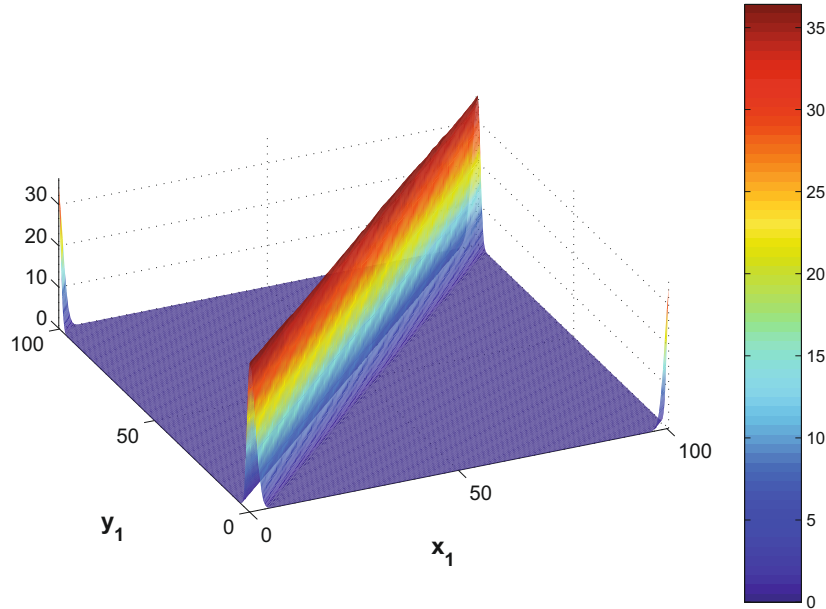
FIGURE 6. Computed values of the diffusion matrix using the same configurations from the O_2-simulation that gave the reaction term in Figure 5(a).

where $k_0 = (r_i q_i)^{-1}$, $k_1 = \gamma q_i^{-1}$, $\gamma = k_B T$ and the reaction term $\overline{A_0}(\bar{m}; O_1)$ is independent of the orientation. Figure 6 shows the computed diffusion coefficients for the noise with the O_2 orientation.

5. THE COMPUTATIONAL METHOD

This section describes the numerical implementation for computing the drift and diffusion functions, defining the coarse-grained phase-field dynamics (4.1), from two-phase Smoluchowski dynamics simulations at constant volume and temperature. The procedure computes time averaged quantities like the time averaged potential energy phase-field

$$\lim_{\mathcal{T} \to \infty} \mathcal{T}^{-1} \mathbb{E}\left[\int_0^{\mathcal{T}} m^t \mathrm{d}t \right]$$

and the corresponding coarse-grained drift and diffusion coefficient functions (4.11) and (4.10). A more detailed description of the numerical method and the computational results is in [29].

5.1. Discrete Smoluchowski system simulated at constant volume

The discrete time approximations $\bar{X}^n$ of X^{t_n} were computed using the explicit Euler-Maruyama scheme

$$\bar{X}^n = \bar{X}^{n-1} - \partial_X V(\bar{X}^{n-1}) \, \Delta t^n + \sqrt{2 k_B T} \, \Delta W^n, \tag{5.1}$$

where $\Delta t^n = t^n - t^{n-1}$ is a time increment and $\Delta W^n = W(t^n) - W(t^{n-1})$ is an increment in the $3N$-dimensional Wiener process. Each run was performed using constant time step size.

The particles are contained in a computational cell, shaped like a rectangular box, of fixed dimensions and the boundary conditions are periodic in all directions. Hence the volume, V, and the number of particles, N,

are fixed. In standard Hamiltonian molecular dynamics the temperature is determined by the kinetic energy due to velocity fluctuations; in Smoluchowski dynamics the temperature, T, enters directly in the dynamics. The temperature parameter is held fixed, which can be viewed as a kind of thermostat built into the dynamics simulating the canonical ensemble with constant number of particles, volume, and temperature. Since the volume of the computational cell is constant the overall density of the system remains constant over time, which allows for stationary two-phase configurations where part of the domain is solid and part is liquid.

The initial configurations for the Smoluchowski dynamics simulations were set up to obtain a two-phase system at temperature $T = 2.90$ with approximately equal volumes of solid and liquid and with stationary interfaces; see Figure 2. Simulation O_1 used 64 131 particles in a computational cell of dimensions $93.17 \times 23.29 \times 23.29$, while simulation O_2 used 78 911 particles in a cell of dimensions $100.86 \times 24.71 \times 24.96$. The values above are given in reduced Lennard-Jones units.

An effect of the finite size of the computational cell is that periodic boundary conditions may interact with the solid and affect the results; here the computational cell was chosen to match the FCC structure in a specific orientation with respect to the box and thus stabilised the structure and orientation. It is important to know that the density in the FCC part (and hence the box cross section) is consistent with constant pressure simulations close to the melting point. A related question is whether the length of the computational box is large enough for properties around the interfaces in the infinitely layered structure to be good approximations of those near an interface between a solid and liquid on the macroscopic scale.

5.2. Computation of the coarse-grained model functions

The drift and diffusion coefficient functions in the stochastic differential equation (4.1) for the coarse-grained phase-field are defined in terms of the time averaged expected values (4.11) and (4.10) on the form

$$\frac{1}{\mathcal{T}}\mathbb{E}\left[\int_0^{\mathcal{T}} \psi(\cdot; X^t) \,\Big|\, X^0 = X_0\right],$$

where the initial configuration X_0 is a configuration of a stationary two-phase system. By setting up X_0 and simulating discrete sample trajectories using the Euler-Maruyama method (5.1), a sequence of configurations $\{\bar{X}^k\}_{k=1}^K$ approximating the sequence $\{X^{t_k}\}_{k=1}^K$ for some times $0 < t_1 < \ldots < t_K = \mathcal{T}$ is obtained. In a post processing step a set of configurations $\mathcal{S} \subseteq \{\bar{X}^k\}_{k=1}^K$ is selected and averages

$$\mathcal{A}_{\mathcal{S}}\left(\psi\right) = \sum_{X \in \mathcal{S}} \psi(\cdot; X) w_X,$$

consistently weighted with weights w_X, are computed as approximations of the corresponding expected values in the continuous time model. It is usually more efficient not to include every configuration in the averages; see [29].

As described in Section 4.3, the averages are functions of the coordinate direction x_1, normal to the planar interface, since the mollifier in the definition (2.4) of the microscale phase-field, m, is chosen to take uniform averages in the planes parallel to the interface. The mollifier used in the computations is

$$\eta(x) = \eta(x_1) = c \exp\left(-\frac{1}{2}\left(\frac{x_1}{\epsilon}\right)^2\right) \mathbf{1}_{|x_1| < R_c}, \tag{5.2}$$

where c is a normalising constant, ϵ is a smoothing parameter, and R_c is a cut-off. The smoothing parameter is on the order of typical nearest neighbour distances, $\epsilon = 1$, and $R_c = 6\epsilon$ which gives $\eta(R_c) \approx 1.5 \times 10^{-8}\eta(0)$.

An explicit derivation of expressions for the drift and the diffusion is given in the appendix of [29]. Separating the drift in terms containing two, one, and zero derivatives of the mollifier, the right hand side of (4.11) is

approximated by

$$k_{\mathrm{B}}T\frac{\partial^2}{\partial x_1^2}\mathcal{A}_{\mathcal{S}}\left(m\right)+\frac{\partial}{\partial x_1}\mathcal{A}_{\mathcal{S}}\left(a_1\right)+\mathcal{A}_{\mathcal{S}}\left(a_0\right),$$

where

$$a_1(x;X)=\sum_{j=1}^{N}(k_{\mathrm{B}}T-m_j(X))[F_j(X)]_1\eta(x-X_j) \tag{5.3}$$

and

$$a_0(x;X)=-\sum_{j=1}^{N}\left(k_{\mathrm{B}}T\partial_{X_j}\cdot F_j(X)+\frac{1}{2}\|F_j(X)\|^2\right)\eta(x-X_j)-\frac{1}{2}\sum_{j=1}^{N}\sum_{i\neq j,i=1}^{N}f_{ij}(X)\cdot F_j(X)\eta(x-X_i). \tag{5.4}$$

Here F_j is the total force acting on particle j, $[F_j(X)]_1$ is the x_1-component of the force, and f_{ij} are the contributions from individual pairs,

$$F_j(X)=-\partial_{X_j}V(X)=\sum_{i\neq j,i=1}^{N}\Phi'(\|X_i-X_j\|)\frac{X_i-X_j}{\|X_i-X_j\|}=\sum_{i\neq j,i=1}^{N}f_{ij}(X).$$

The right hand side in equation (4.10), for the coarse-grained diffusion, is approximated by

$$\overline{B}(\cdot,\cdot)=\mathcal{A}_{\mathcal{S}}\left(2k_{\mathrm{B}}T\sum_{j=1}^{N}\left(p_j(\cdot,\cdot;X)+q_j(\cdot,\cdot;X)\right)\right), \tag{5.5}$$

where

$$p_j(x,y;X)=\left(\frac{m_j(X)}{\epsilon^2}\right)^2\left[x-X_j\right]_1\left[y-X_j\right]_1\eta(x-X_j)\eta(y-X_j)$$
$$-\frac{m_j(X)}{2\epsilon^2}[x-X_j]_1\eta(x-X_j)\left([F_j(X)]_1\eta(y-X_j)+\sum_{i\neq j,i=1}^{N}[f_{ij}(X)]_1\eta(y-X_i)\right)$$
$$-\frac{m_j(X)}{2\epsilon^2}[y-X_j]_1\eta(y-X_j)\left([F_j(X)]_1\eta(x-X_j)+\sum_{i\neq j,i=1}^{N}[f_{ij}(X)]_1\eta(x-X_i)\right)$$

and

$$q_j(x,y;X)=\frac{1}{4}\left(F_j(X)\eta(x-X_j)+\sum_{i\neq j,i=1}^{N}f_{ij}(X)\eta(x-X_i)\right)\cdot\left(F_j(X)\eta(y-X_j)+\sum_{i\neq j,i=1}^{N}f_{ij}(X)\eta(y-X_i)\right).$$

The functions $\mathcal{A}_{\mathcal{S}}\left(\psi\right)$ are computed in a discrete set of points $D_K=\{x_1^i\}_{i=1}^K$ along the x_1 axis of the molecular dynamics domain. This makes the computed components, $\mathcal{A}_{\mathcal{S}}\left(m\right)$, $\mathcal{A}_{\mathcal{S}}\left(a_1\right)$, and $\mathcal{A}_{\mathcal{S}}\left(a_0\right)$, of the drift coefficient function K-vectors and the computed $\overline{B}$ a K-by-K matrix. The individual diffusion coefficient functions $\overline{b}_j$ are obtained by taking the square root of the computed diffusion matrix, $\overline{B}=\overline{B}^{1/2}(\overline{B}^{1/2})^{\mathrm{T}}$, and letting the j-th column of $\overline{B}^{1/2}$ define $\overline{b}_j$.

With one Wiener process $\widetilde{W}_j$ in the coarse-grained stochastic differential equation (4.1) per evaluation point, $K = M$, the component vectors, $\overline{b}_j$, of the diffusion in the coarse-grained equation can be defined as the column vectors of the matrix B, to obtain

$$\sum_{j=1}^{M} \overline{b}_j \overline{b}_j^{\mathrm{T}} \approx \overline{B}.$$

If two grid points, x_1 and y_1, are further apart than twice the sum of the cut-off in the potential and the cut-off in the mollifier, then $p_j(x, y; \cdot)$ and $q_j(x, y; \cdot)$ are zero; hence a natural ordering $x_1^1 < x_1^2 < \ldots < x_1^K$ of the grid points makes $\overline{B}$ a band matrix. Using the square root of $\overline{B}$ is a way that preserves the connection between grid points and diffusion functions and hence the dominating terms in a tabulated vector $\overline{b}_j$ are those of nearby grid points.

Remark 5.1 (extensions)**.** We have computed coarse-grained approximations for a pair potential molecular equilibrium system in the canonical ensemble (with constant number of particles, volume and temperature) using periodic boundary conditions.

In macroscopic phase transitions the system often has constant pressure instead of volume. Therefore it is interesting to extend the work to molecular systems with constant pressure, by controlling the pressure, *cf.* [9]. At constant pressure, the enthalpy, which is the sum of the internal energy and the pressure times the volume, pv, takes the role the internal energy has at constant volume, so that at constant pressure the potential V here is essentially replaced by $V + pv$ and a varying computational volume with pressure boundary conditions.

It would also be interesting to simulate a molecular system above or below the melting temperature. The situation with a stationary phase transition at the melting temperature then changes to a moving interface. To follow a traveling wave leads to molecular dynamics with non periodic boundary conditions. To simulate molecular dynamics in any situation different from the case with periodic boundary conditions is non standard and much remains to understand, *cf.* [8]. Our setting with a phase transition in a Lennard-Jones type system would be suitable for learning about other boundary conditions: on one side the system is crystallized, with a certain orientation, and on the opposite side the system is liquid. In a coordinate system, moving with the propagation speed v of the traveling wave, the molecular system has a stationary phase interface, as at the melting temperature. The translated Smoluchowski dynamics

$$\mathrm{d}X^t = -\left(\partial_X V(X^t) - v\right)\mathrm{d}t + (2T)^{1/2}\mathrm{d}W^t$$

could then be used to find the correct propagation speed v, as the speed where the translated molecular dynamics has a stationary phase transition. The velocity v could be used to validate the method. The other new ingredient here is to handle particles leaving the domain through the two surfaces with normals in the direction of the wave propagation. The remaining identification of the traveling wave and the coarse-graining then becomes the same as for the case at the melting temperature.

Remark 5.2 (hybrid method)**.** There may be cases when the phase-field method cannot resolve a local phenomenon in a small domain D. A more accurate solution could then be to couple the phase-field method to a molecular dynamics simulation in the domain D. Since the coarse-grained potential energy is identified as the phase-field in the domain outside D, the natural condition for the phase-field at the boundary ∂D is to set the phase-field variable equal to the coarse-grained potential energy determined from molecular dynamics in D. Then the molecular dynamics periodic boundary conditions have to be replaced with some conditions coupled to the macroscopic phase-field and temperature.

Remark 5.3 (variances)**.** Note that the approximation error $\mathbb{E}\big[g\big(m(X^{\mathcal{T}},\cdot)\big) - g(\bar{m}^{\mathcal{T}})\big]$ becomes proportional to the variances

$$\mathbb{E}\bigg[\int_0^{\mathcal{T}} \langle \alpha - a, \alpha - a\rangle \mathrm{d}t\bigg],$$

$$\mathbb{E}\bigg[\int_0^{\mathcal{T}} \bigg\langle \sum_{j=1}^N \beta_j \otimes \beta_j - \sum_{k=1}^M b_k \otimes b_k, \sum_{j=1}^N \beta_j \otimes \beta_j - \sum_{k=1}^M b_k \otimes b_k \bigg\rangle \mathrm{d}t\bigg].$$

The first variations $\partial\bar{u}'\big(m(X^t,\cdot),t\big)/\partial\alpha$ and $\partial\bar{u}''\big(m(X^t,\cdot),t\big)/\partial\beta_j$ determine the factors of proportionality.

6. Conclusions

This report presents a method to computationally determine a phase-field equation, related to previous work on atomistic simulations following traveling waves, *cf.* [12]. The new ingredient here is that the whole reaction function (not only a parameter) and the noise function are determined, for a special case of a stationary wave corresponding to a simulation at the melting temperature.

The method to determine the stochastic phase-field model has three steps:

- a precise quantitative atomistic definition of the phase-field variable $m : \mathbb{R}^{3N} \times \mathbb{R}^3 \to \mathbb{R}$, based on a given molecular dynamics potential energy $V : \mathbb{R}^{3N} \to \mathbb{R}$ in (2.4);
- derivation of a macroscopic coarse-grained dynamics, *i.e.* the phase-field model (4.13)

$$(rq)^{-1}\mathrm{d}\bar{m} = \big(k_{\mathrm{B}}Tq^{-1}\partial_x^2\bar{m} + \overline{A_0}(\bar{m})\big)\mathrm{d}t + r^{-1/2}q^{-1}\sum_k \bar{b}_k(\bar{m})\mathrm{d}\tilde{W}_k^t,$$

  for $\bar{m}(x,t)$ approximating $m(X^t,x)$ in (4.1)–(4.4), using the microscopic Smoluchowski molecular dynamics

$$\mathrm{d}X_i^t = -\partial_{X_i}V(X^t)\mathrm{d}t + \sqrt{2k_{\mathrm{B}}T}\,\mathrm{d}W_i^t,$$

  for particle positions $X : [0,\infty) \to \mathbb{R}^{3N}$ and standard Brownian motions $\tilde{W}_k$ and W_i;
- numerical molecular dynamics computation to evaluate m and $\bar{b}$ in the phase transition by ensemble averages based on (4.4); the phase-field variable m then determines the functions $\bar{A}_0$ and q from a traveling wave: the shape of the traveling wave, computed from molecular dynamics, determines the reaction term balancing the diffusion (since the total drift becomes zero in our case of a stationary wave), the reaction and diffusion terms in the phase-field equation are then determined up to a multiplicative constant – this multiplicative constant is chosen to match equilibrium fluctuations of the phase-field; the diffusion coefficient $\bar{b}$ is determined from an optimization problem, where the diffusion is chosen to minimize the error of observables in the Gibbs ensemble.

The coarse-grained model approximates Gibbs ensemble averages of the atomistic phase-field, by choosing coarse-grained drift and diffusion functions that minimize the approximation error of observables in this ensemble average.

The quotient of reaction terms $q = q_i := \overline{A_0(O_i)}/\overline{A_0(O_1)}$ depends on the crystal orientation O_i relative to the phase interface and the reaction term $\bar{A}_0$ forms a double well potential shown in Figure 5. The diffusion coefficient $\bar{b}$ is determined from the square root of the diffusion matrix of the noise in Figure 6. The function $r = r_i$ is the quotient of relaxation times: the time scale τ used Smoluchowski dynamics may depend on the crystal orientation, since the relaxation time τ may differ for two orientations O_i, and we define

$r_i := \tau(O_i)/\tau(O_1)$ for the reference orientation O_1. Langevin dynamics

$$\mathrm{d}X_i^s = v_i^s \, \mathrm{d}s,$$

$$M\mathrm{d}v_i^s = -\partial_{X_i}V(X^s) \, \mathrm{d}s - \frac{v_i^s}{\tau} \, \mathrm{d}s + \sqrt{\frac{2\gamma}{\tau}} \, \mathrm{d}\hat{W}_i^s,$$

defines the relaxation time. The relaxation time τ is related to the shear viscosity η by $\tau(O_i)/\tau(O_1) = \eta(O_1)/\eta(O_i)$. The function r can therefore be determined similarly as the function $m(x)$, replacing the Smoluchowski dynamics with Langevin dynamics (or Hamiltonian dynamics corresponding to $\tau = \infty$) and using a standard molecular dynamics expression for the viscosity, *cf.* [9].

The computational results with a Lennard-Jones type molecular potential shows that the determined phase-field equation has the right qualitative behavior, with a double well potential and orientation dependent coefficients. Further numerical studies would be needed to validate the method for realistic materials.

References

[1] G. Amberg, Semi sharp phase-field method for quantitative phase change simulations. *Phys. Rev. Lett.* **91** (2003) 265505–265509.

[2] M. Asta, C. Beckermann, A. Karma, W. Kurz, R. Napolitano, M. Plapp, G. Purdy, M. Rappaz and R. Trivedi, Solidification microstructures and solid-state parallels: Recent developments, future directions. *Acta Mater.* **57** (2009) 941–971.

[3] J.T. Beale and A.J. Majda, Vortex methods. I. Convergence in three dimensions. *Math. Comp.* **39** (1982) 1–27.

[4] W.J. Boettinger, J.A. Warren, C. Beckermann and A. Karma, Phase field simulation of solidification. *Ann. Rev. Mater. Res.* **32** (2002) 163–194.

[5] E. Burman and J. Rappaz, Existence of solutions to an anisotropic phase-field model. *Math. Methods Appl. Sci.* **26** (2003) 1137–1160.

[6] E. Cances, F. Legoll and G. Stolz, Theoretical and numerical comparison of some sampling methods for molecular dynamics. *ESAIM: M2AN* **41** (2007) 351–389.

[7] A. De Masi, E. Orlandi, E. Presutti and L. Triolo, Glauber evolution with the Kac potentials. I. Mesoscopic and macroscopic limits, interface dynamics. *Nonlinearity* **7** (1994) 633–696.

[8] W. E and W. Ren, Heterogeneous multiscale method for the modeling of complex fluids and micro-fluidics. *J. Comput. Phys.* **204** (2005) 1–26.

[9] D. Frenkel and B. Smit, *Understanding Molecular Simulation.* Academic Press (2002).

[10] J. Gooodman, K.-S. Moon, A. Szepessy, R. Tempone and G. Zouraris, *Stochastic Differential Equations: Models and Numerics.* http://www.math.kth.se/$\sim$szepessy/sdepde.pdf.

[11] R.J. Hardy, Formulas for determining local properties in molecular dynamics: shock waves. *J. Chem. Phys.* **76** (1982) 622–628.

[12] J.J. Hoyt, M. Asta and A. Karma, Atomistic and continuum modeling of dendritic solidification. *Mat. Sci. Eng. R* **41** (2003) 121–163.

[13] J.H. Irving and J.G. Kirkwood, The statistical mechanical theory of transport processes. IV. The equations of hydrodynamics. *J. Chem. Phys.* **18** (1950) 817–829.

[14] L.P. Kadanoff, *Statistical physics: statics, dynamics and renormalization.* World Scientific (2000).

[15] A. Karma and W.J. Rappel, Phase-field model of dendritic side branching with thermal noise. *Phys Rev. E* **60** (1999) 3614–3625.

[16] M. Katsoulakis and A. Szepessy, Stochastic hydrodynamical limits of particle systems. *Comm. Math. Sciences* **4** (2006) 513–549.

[17] P.R. Kramer and A.J. Majda, Stochastic mode reduction for particle-based simulation methods for complex microfluid systems. *SIAM J. Appl. Math.* **64** (2003) 401–422.

[18] H. Kroemer and C. Kittel, *Thermal Physics.* W.H. Freeman Company (1980).

[19] L.D. Landau and E.M. Lifshitz, *Statistical Physics Part 1.* Pergamon Press (1980).

[20] T.M. Liggett, *Interacting particle systems.* Springer-Verlag, Berlin (2005).

[21] S. Mas-Gallic and P.-A. Raviart, A particle method for first-order symmetric systems. *Numer. Math.* **51** (1987) 323–352.

[22] G. Morandi, F. Napoli and E. Ercolessi, *Statistical Mechanics: An Intermediate Course.* World Scientific Publishing (2001).

[23] E. Nelson, *Dynamical Theories of Brownian Motion.* Princeton University Press (1967).

[24] S.G. Pavlik and R.J. Sekerka, Forces due to fluctuations in the anisotropic phase-field model of solidification. *Physica A* **268** (1999) 283–290.

[25] T. Schlick, *Molecular modeling and simulation*. Springer-Verlag (2002).

[26] H.M. Soner, Convergence of the phase-field equations to the Mullins-Sekerka Problem with kinetic undercooling. *Arch. Rat. Mech. Anal.* **131** (1995) 139–197.

[27] A. Szepessy, R. Tempone and G. Zouraris, Adaptive weak approximation of stochastic differential equations. *Comm. Pure Appl. Math.* **54** (2001) 1169–1214.

[28] N.G. van Kampen, *Stochastic Processes in Physics and Chemistry*. North-Holland (1981).

[29] E. von Schwerin, *A Stochastic Phase-Field Model Computed From Coarse-Grained Molecular Dynamics*. `arXiv:0908.1367`, included in [30].

[30] E. von Schwerin, *Adaptivity for Stochastic and Partial Differential Equations with Applications th Phase Transformations*. Ph.D. Thesis, KTH, Royal Institute of Technology, Stockholm, Sweden (2007).