Estimation of a Gas Diffusion Coefficient by Fitting Molecular Dynamics Trajectories to Finite-Difference Simulations

Isaac Viviano[†]

Project advisor: Robert Krasny[‡]

Abstract. A procedure is presented to estimate the diffusion coefficient of a uniform patch of argon gas in a uniform background of helium gas. Molecular Dynamics (MD) simulations of the two gases interacting through the Lennard-Jones potential are carried out using the LAMMPS software package. In addition, finite-difference (FD) calculations are used to solve the continuum diffusion equation for the argon concentration with a given diffusion coefficient. To contain the computational cost and facilitate data visualization, both MD and FD computations were done in two space dimensions. The MD argon trajectories were binned to the FD grid, and the optimal diffusion coefficient was estimated by minimizing the difference between the binned MD data and the FD solution with a nonlinear least squares procedure (Levenberg–Marquardt algorithm). Numerical results show the effect of the MD binning parameter and FD grid spacing. The estimated diffusion coefficient is compared to an experimental measurement.

Key words. gas diffusion, diffusion equation, finite-difference method, molecular dynamics, LAMMPS, parameter estimation

1. Introduction.

The diffusion of a solute gas (for example argon) through a solvent gas (for example helium) can be quantified by a diffusion coefficient D. The coefficient can be measured experimentally by gas chromatography, which separates a mixture of gases into pure components as it travels through a column so that D can be calculated from the peak widths of the mass spectrum [12]. The diffusion coefficient can also be calculated theoretically from the expression

$$D = \frac{1}{6} \lim_{t\to\infty} \left\langle \frac{\mathrm{d}}{\mathrm{d}t} (\mathbf{r}(t) - \mathbf{r}(0))^2 \right\rangle, \tag{1.1}$$

where $\mathbf{r}(t)$ is the trajectory of a gas particle and $\langle\cdot\rangle$ represents the average over all diffusing particles [5]. The derivative of the square displacement describes how fast each particle spreads out, so for sufficiently large simulation times, the ensemble average in (1.1) quantifies the overall diffusion rate. One way to generate the particle trajectories for (1.1) is with Molecular Dynamics (MD) simulations, which provide the trajectories of a collection of particles evolving by Hamiltonian dynamics in a classical interparticle force field. Here, we present an alternative theoretical approach to (1.1) for determining the gas diffusion coefficient.

Fick's laws give a continuum description of diffusion, resulting in a partial differential equation for the time-evolution of the solute concentration [2]. Our approach estimates the the diffusion coefficient D by minimizing the discrepancy between numerical finite-difference (FD) solutions of the diffusion equation and MD simulations of the solute concentration. We anticipate that this procedure can be applied to estimate coefficients in a differential

[†](iviviano@oberlin.edu)

[‡]University of Michigan, Ann Arbor, MI

equation for other physical systems described by a continuum model such as phase separation of polymers.

Our MD simulations are carried out using LAMMPS (Large-scale Atomic/Molecular Massively Parallel Simulator), an open-source classical MD package focused on materials modeling, which was developed by researchers at Sandia National Laboratories and collaborators [9]. We fix values of the state variables (pressure and temperature) and conduct equilibrium MD simulations where the gas particles follow Hamiltonian dynamics of pairwise Lennard-Jones interactions. Since MD simulations can incorporate the effect of varying the state variables, our procedure could be applied to estimate the dependence of the diffusion coefficient on these variables.

In general, MD and FD calculations may be applied to one-, two-, and three-dimensional systems, and throughout this work the spatial dimension is denoted d. Similarly, LAMMPS provides the functionality to conduct MD simulations in one-, two-, or three-dimensions. While the models are equally valid for any dimension, the estimation procedure is implemented for $d = 2$ to contain the computational cost and facilitate data visualization.

The rest of the article is organized as follows. The derivation of the diffusion equation is given in section 2, and the FD methods are presented in section 3. The MD approach is described in section 4, and the MD and FD simulation results are compared in section 5. The diffusion coefficient estimation algorithm is presented in section 6, and the estimates are given in section 7. The work is summarized in section 8.

2. Diffusion equation.

This section introduces the continuum model for gas diffusion, explaining its physical basis and its solution by Fourier series. The gas concentration is a scalar function of space and time, $u(\mathbf{x}, t)$, $\mathbf{x} \in \mathbb{R}^d$. For a system of gas particles subject to a non-zero concentration gradient, random motion tends to reduce the gradient over time. The net movement of particles from regions of higher concentration to lower concentration is called diffusion and these concepts are quantified by Fick's laws. While our numerical methods do not make use of the analytical solution, it prescribes important physical properties of the differential equation that the finite-difference methods should replicate. In particular, the periodic boundary conditions ensure that mass is conserved and energy decreases in time.

2.1. Derivation from Fick's laws.

Fick's first law says that the diffusion flux $\mathbf{J}(\mathbf{x}, t)$ is proportional to the concentration gradient $\nabla u(\mathbf{x}, t)$ with proportionality constant given by a scalar diffusion coefficient D,

$$\mathbf{J}(\mathbf{x}, t) = -D\nabla u(\mathbf{x}, t). \tag{2.1}$$

In general the diffusion coefficient depends on many factors such as the gas temperature and pressure, but in this work it is assumed to be a given constant. Conservation of mass states that the rate of change of the mass contained in a volume V is determined by the mass flux through the surface $S = \partial V$ bounding the volume V,

$$\frac{d}{dt}\int_V u(\mathbf{x}, t)\mathrm{d}\mathbf{x} + \oint_S \mathbf{J}(\mathbf{x}, t) \cdot \mathrm{d}\mathbf{S} = 0. \tag{2.2}$$

Application of the divergence theorem converts the surface integral into a volume integral which yields the differential form of mass conservation,

$$\frac{\partial}{\partial t}u(\mathbf{x},t) + \nabla \cdot \mathbf{J}(\mathbf{x},t) = 0. \tag{2.3}$$

Then substituting (2.1) into (2.3) yields Fick's second law,

$$\frac{\partial}{\partial t}u(\mathbf{x},t) = D\nabla^2 u(\mathbf{x},t), \tag{2.4}$$

a partial differential equation for the time-evolution of the gas concentration known as the diffusion equation [2]. The operator $\nabla^2 = \nabla\cdot\nabla$ on the right side of (2.4) is the Laplacian. This work considers the diffusion equation with periodic boundary conditions on the d-dimensional unit cube $\Omega = [0,1]^d$ and a given initial condition,

$$u(\mathbf{x},t)\big|_{x_i=0} = u(\mathbf{x},t)\big|_{x_i=1}, \quad \frac{\partial}{\partial x_i}u(\mathbf{x},t)\big|_{x_i=0} = \frac{\partial}{\partial x_i}u(\mathbf{x},t)\big|_{x_i=1}, \quad i = 1,\ldots,d, \tag{2.5a}$$

$$u(\mathbf{x},0) = u_0(\mathbf{x}). \tag{2.5b}$$

Note that (2.4) is also known as the heat equation since it models heat flow when u represents the material temperature and D is its thermal conductivity. The next subsection derives an analytic expression for the solution of the diffusion equation.

2.2. Analytical solution by Fourier Series. For the d-dimensional unit cube with periodic boundary conditions, the solution of the diffusion equation (2.4) can be expressed as a Fourier series [7],

$$u(\mathbf{x},t) = \sum_{\mathbf{m}\in\mathbb{Z}^d} \widehat{u}_{\mathbf{m}}(t)\mathrm{e}^{2\pi\mathrm{i}\mathbf{m}\cdot\mathbf{x}}, \quad \widehat{u}_{\mathbf{m}}(t) = \int_\Omega u(\mathbf{x},t)\ \mathrm{e}^{-2\pi\mathrm{i}\mathbf{m}\cdot\mathbf{x}}\ \mathrm{d}\mathbf{x}, \tag{2.6}$$

where $\widehat{u}_{\mathbf{m}}(t)$ denotes the $\mathbf{m}$-th Fourier coefficient of $u(\mathbf{x},t)$ with wavenumbers over the d-dimensional integer lattice $\mathbf{m}\in\mathbb{Z}^d$. From substituting (2.6) into (2.4) we obtain the initial value problem,

$$\frac{\mathrm{d}}{\mathrm{d}t}\widehat{u}_{\mathbf{m}}(t) = -4\pi^2|\mathbf{m}|^2 D\widehat{u}_{\mathbf{m}}(t), \quad \widehat{u}_{\mathbf{m}}(0) = \int_\Omega u_0(\mathbf{x})\ \mathrm{e}^{-2\pi\mathrm{i}\mathbf{m}\cdot\mathbf{x}}\ \mathrm{d}\mathbf{x}, \tag{2.7}$$

for the Fourier coefficients. Its solution yields the time dependence,

$$\widehat{u}_{\mathbf{m}}(t) = \widehat{u}_{\mathbf{m}}(0)\mathrm{e}^{-4\pi^2|\mathbf{m}|^2 Dt}, \tag{2.8}$$

where $\widehat{u}_{\mathbf{m}}(0)$ is the $\mathbf{m}$-th Fourier coefficient of the initial condition u_0. Several properties of the solution can now be derived.

First, setting $\mathbf{m} = \mathbf{0}$ in (2.8) yields $\widehat{u}_{\mathbf{0}}(t) = \widehat{u}_{\mathbf{0}}(0)$, which implies that the total mass of the system is conserved in time,

$$\int_\Omega u(\mathbf{x},t)\ \mathrm{d}\mathbf{x} = \int_\Omega u_0(\mathbf{x})\ \mathrm{d}\mathbf{x}. \tag{2.9}$$

In addition, non-constant modes $\mathbf{m} \neq \mathbf{0}$ decay exponentially in time, and hence the concentration $u(\mathbf{x}, t)$ converges to its average value $\int_\Omega u_0(\mathbf{x})\, \mathrm{d}\mathbf{x}$ as $t \to \infty$.

Another important property of solutions to the diffusion equation with periodic boundary conditions is that their energy is non-increasing in time. The energy is given by the square of the 2-norm of the solution,

$$E(t) = \|u(\cdot, t)\|_2^2 = \int_\Omega |u(\mathbf{x}, t)|^2 \mathrm{d}\mathbf{x} = \sum_{\mathbf{m} \in \mathbb{Z}^d} |\widehat{u}_{\mathbf{m}}(t)|^2, \tag{2.10}$$

where the last equality follows from Parseval's relation which states that the 2-norm of a periodic function equals the discrete 2-norm of its Fourier coefficients. Using the ODE for the Fourier coefficients (2.7), it follows that

$$E'(t) = 2 \sum_{\mathbf{m} \in \mathbb{Z}^d} \widehat{u}_{\mathbf{m}}(t) \frac{\mathrm{d}}{\mathrm{d}t} \widehat{u}_{\mathbf{m}}(t) = -8\pi^2 D \sum_{\mathbf{m} \in \mathbb{Z}^d} |\mathbf{m}|^2 |\widehat{u}_{\mathbf{m}}(t)|^2 \leq 0, \tag{2.11}$$

which implies that the energy is non-increasing in time. The next section discusses finite-difference methods which are used to solve the diffusion equation numerically.

3. Finite-Difference method.

In a finite-difference method the temporal and spatial domains are discretized and the derivatives in the equation are replaced by finite-difference approximations that enable the numerical solution to be found by stepping forward in time. This work considers two such methods for solving the diffusion equation (2.4) with periodic boundary conditions and initial condition (2.5). A time step k and space step $h = 1/N$ are chosen, where N is a positive integer. The associated spatial grid is $\Omega_h = h[N]^d$, where $[N] = \{0, \ldots, N-1\}$, and the solution $u(\mathbf{x}, t)$ is approximated at grid points $\mathbf{x}_\mathbf{j} = h\mathbf{j}$ and time steps $t_n = nk$,

$$u_\mathbf{j}^n \approx u(\mathbf{x}_\mathbf{j}, t_n), \quad \mathbf{j} = (j_1, \ldots, j_d) \in [N]^d, \quad n = 0, \ldots, n_{\max}. \tag{3.1}$$

Next, we consider the spatial discretization of the Laplacian.

3.1. Central difference approximation in space.

This work employs a central difference approximation for the spatial discretization of the Laplace operator, yielding the discrete Laplacian,

$$\nabla_h^2 u_\mathbf{j}^n = \frac{1}{h^2} \sum_{i=1}^{d} (u_{\mathbf{j}+\mathbf{e}_i}^n - 2u_\mathbf{j}^n + u_{\mathbf{j}-\mathbf{e}_i}^n), \quad \mathbf{j} \in [N]^d, \tag{3.2}$$

where $\mathbf{e}_i$ is the i-th standard basis vector in $\mathbb{R}^d$. The vector $u_\mathbf{j}^n$ is extended to an N-periodic signal, $u_\mathbf{j}^n = u_{\mathbf{j}+N\mathbf{e}_i}^n$ for $\mathbf{j} \in \mathbb{Z}^d, i = 1, \ldots, d$. This justifies the discrete Fourier series expansion,

$$u_\mathbf{j}^n = \frac{1}{N^d} \sum_{\mathbf{m} \in [N]^d} \widehat{u}_\mathbf{m}^n \mathrm{e}^{2\pi \mathrm{i} \mathbf{m} \cdot \mathbf{j}/N}, \quad \widehat{u}_\mathbf{m}^n = \sum_{\mathbf{j} \in [N]^d} u_\mathbf{j}^n \mathrm{e}^{-2\pi \mathrm{i} \mathbf{m} \cdot \mathbf{j}/N}, \tag{3.3}$$

where $\widehat{u}^n_{\mathbf{m}}$ is the $\mathbf{m}$-th discrete Fourier coefficient of the numerical solution $u^n_{\mathbf{j}}$. Using properties of the exponential function, we can understand how the discrete Laplacian acts in Fourier space,

$$\begin{aligned}
\nabla_h^2 u^n_{\mathbf{j}} &= \frac{1}{h^2}\sum_{i=1}^{d}\frac{1}{N^d}\sum_{\mathbf{m}\in[N]^d}\widehat{u}^n_{\mathbf{m}}(\mathrm{e}^{2\pi\mathrm{i}\mathbf{m}\cdot(\mathbf{j}+\mathbf{e}_i)/N} - 2\mathrm{e}^{2\pi\mathrm{i}\mathbf{m}\cdot\mathbf{j}/N} + \mathrm{e}^{2\pi\mathrm{i}\mathbf{m}\cdot(\mathbf{j}-\mathbf{e}_i)/N}) \\
&= \frac{1}{N^d}\sum_{\mathbf{m}\in[N]^d}\sum_{i=1}^{d}\widehat{u}^n_{\mathbf{m}}\frac{1}{h^2}(\mathrm{e}^{2\pi\mathrm{i}m_i/N} + \mathrm{e}^{-2\pi\mathrm{i}m_i/N} - 2)\mathrm{e}^{2\pi\mathrm{i}\mathbf{m}\cdot\mathbf{j}/N} \\
&= \frac{1}{N^d}\sum_{\mathbf{m}\in[N]^d}\widehat{u}^n_{\mathbf{m}}\frac{1}{h^2}\sum_{i=1}^{d}(\mathrm{e}^{2\pi\mathrm{i}m_i/N} + \mathrm{e}^{-2\pi\mathrm{i}m_i/N} - 2)\mathrm{e}^{2\pi\mathrm{i}\mathbf{m}\cdot\mathbf{j}/N} \\
&= \sum_{\mathbf{m}\in[N]^d}\lambda_{\mathbf{m}}\widehat{u}^n_{\mathbf{m}}\mathrm{e}^{2\pi\mathrm{i}\mathbf{m}\cdot\mathbf{j}/N}, \quad \lambda_{\mathbf{m}} = \frac{1}{N^d}\frac{1}{h^2}\sum_{i=1}^{d}(\mathrm{e}^{2\pi\mathrm{i}m_i/N} + \mathrm{e}^{-2\pi\mathrm{i}m_i/N} - 2),
\end{aligned} \tag{3.4}$$

where $\lambda_{\mathbf{m}}$ are the eigenvalues and $(\mathrm{e}^{2\pi\mathrm{i}\mathbf{m}\cdot\mathbf{j}/N} : \mathbf{j} \in [N]^d)$ are the corresponding eigenvectors of the discrete Laplacian with periodic boundary conditions. Equation (3.4) illustrates the principle that applying the spatial operator in physical space is equivalent to multiplication in Fourier space. An alternative expression for the eigenvalues is given by

$$\lambda_{\mathbf{m}} = \frac{2}{h^2}\sum_{i=1}^{d}(\cos(2\pi m_i/N) - 1) = -\frac{4}{h^2}\sum_{i=1}^{d}\sin^2(\pi m_i/N), \quad \mathbf{m} \in [N]^d, \tag{3.5}$$

and since $0 \le \sin^2(\pi m_i/N) \le 1$, the eigenvalues are nonpositive and bounded,

$$-4d/h^2 \le \lambda_{\mathbf{m}} \le 0. \tag{3.6}$$

Next we consider the stability of the difference scheme. Recall that the energy of the solution of the diffusion equation (2.4) does not increase in time, and the difference scheme should have the same property. This is expressed as the discrete stability condition,

$$\|\mathbf{u}^{n+1}\|_2 \le \|\mathbf{u}^n\|_2, \quad \|\mathbf{u}^n\|_2^2 = \sum_{\mathbf{j}\in[N]^d}|u^n_{\mathbf{j}}|^2, \tag{3.7}$$

where the discrete 2-norm of the numerical solution represents the energy. The discrete version of Parseval's relation relates the 2-norm of the numerical solution at a given time to the 2-norm of its discrete Fourier coefficients by,

$$\|\mathbf{u}^n\|_2^2 = \sum_{\mathbf{j}\in[N]^d}|u^n_{\mathbf{j}}|^2 = \frac{1}{N^d}\sum_{\mathbf{m}\in[N]^d}|\widehat{u}^n_{\mathbf{m}}|^2 = \frac{1}{N^d}\|\widehat{u}^n_{\mathbf{m}}\|_2^2. \tag{3.8}$$

In the following two subsections we consider options for the temporal discretization, and it show that each option yields a recurrence relation for the discrete Fourier coefficients,

$$\widehat{u}^{n+1}_{\mathbf{m}} = \rho_{\mathbf{m}}\widehat{u}^n_{\mathbf{m}}, \tag{3.9}$$

where $\rho_{\mathbf{m}}$ is the amplification factor of the $\mathbf{m}$-th discrete Fourier coefficient. A sufficient condition for the numerical method to satisfy the discrete stability condition (3.7) is

$$|\rho_{\mathbf{m}}| \leq 1 \iff |\widehat{u}_{\mathbf{m}}^{n+1}| \leq |\widehat{u}_{\mathbf{m}}^{n}| \text{ for all wavenumbers } \mathbf{m} \in [N]^d, \tag{3.10}$$

which follows from application of Parseval's relation,

$$\|\mathbf{u}^{n+1}\|_2^2 = \frac{1}{N^d} \sum_{\mathbf{m}\in[N]^d} |\widehat{u}_{\mathbf{m}}^{n+1}|^2 \leq \frac{1}{N^d} \sum_{\mathbf{m}\in[N]^d} |\widehat{u}_{\mathbf{m}}^{n}|^2 = \|\mathbf{u}^{n}\|_2^2. \tag{3.11}$$

3.2. Forward Euler scheme in time. A forward difference in time yields the forward Euler/central difference scheme,

$$\frac{u_{\mathbf{j}}^{n+1} - u_{\mathbf{j}}^{n}}{k} = D\nabla_h^2 u_{\mathbf{j}}^{n}, \tag{3.12}$$

where k is the time step. This can be expressed equivalently in terms of the discrete Fourier coefficients,

$$\widehat{u}_{\mathbf{m}}^{n+1} = \widehat{u}_{\mathbf{m}}^{n} + kD\lambda_{\mathbf{m}}\widehat{u}_{\mathbf{m}}^{n} = \rho_{\mathbf{m}}\widehat{u}_{\mathbf{m}}^{n}, \quad \rho_{\mathbf{m}} = (1 + kD\lambda_{\mathbf{m}}) \text{ for all } \mathbf{m} \in [N]^d, \tag{3.13}$$

where $\rho_{\mathbf{m}}$ is the amplification factor of the scheme. The stability condition (3.10) then gives the following sufficient condition for stability of the scheme,

$$|1 + kD\lambda_{\mathbf{m}}| \leq 1 \text{ for all } \mathbf{m} \in [N]^d. \tag{3.14}$$

Recalling the eigenvalue bound, $-4d/h^2 \leq \lambda_{\mathbf{m}} \leq 0$, in (3.6), the stability condition (3.14) is guaranteed when the time step satisfies

$$|1 + kD\lambda_{\mathbf{m}}| \leq 1 \Leftrightarrow -1 \leq 1 + kD\lambda_{\mathbf{m}} \leq 1 \Leftrightarrow -2 \leq kD\lambda_{\mathbf{m}} \leq 0 \Leftrightarrow k \leq k_c, \tag{3.15}$$

where $k_c = h^2/2dD$ is the critical time step. Note that the lower bound in (3.6) is achieved when $m_i = N/2, i = 1, \ldots, d$ (assuming N is even), so k_c is a tight upper bound on stable time steps. Figure 1 plots the amplification factor of the scheme, $\rho_{\mathbf{m}} = 1 + kD\lambda_{\mathbf{m}}$, for three time step values, $k < k_c, k = k_c, k > k_c$. For simplicity the spatial dimension is $d = 1$. The results indicate stability for $k \leq k_c$ and short wavelength instability for $k > k_c$. For time step values $k > k_c$, the amplification factor lies outside the stable range, $\rho_{\mathbf{m}} < -1$, for high wavenumber modes, indicating a short wavelength instability for the difference scheme in that case. Next we consider a time discretization that does not suffer from this instability.

3.3. Crank-Nicolson scheme. A trapezoidal difference in time yields the Crank-Nicolson scheme,

$$\frac{u_{\mathbf{j}}^{n+1} - u_{\mathbf{j}}^{n}}{k} = \frac{1}{2}D(\nabla_h^2 u_{\mathbf{j}}^{n+1} + \nabla_h^2 u_{\mathbf{j}}^{n}), \tag{3.16}$$

which can be expressed equivalently in terms of the discrete Fourier coefficients of $u_{\mathbf{j}}^{n}$,

$$\widehat{u}_{\mathbf{m}}^{n+1} = \widehat{u}_{\mathbf{m}}^{n} + \frac{1}{2}kD(\lambda_{\mathbf{m}}\widehat{u}_{\mathbf{m}}^{n+1} + \lambda_{\mathbf{m}}\widehat{u}_{\mathbf{m}}^{n}), \tag{3.17}$$

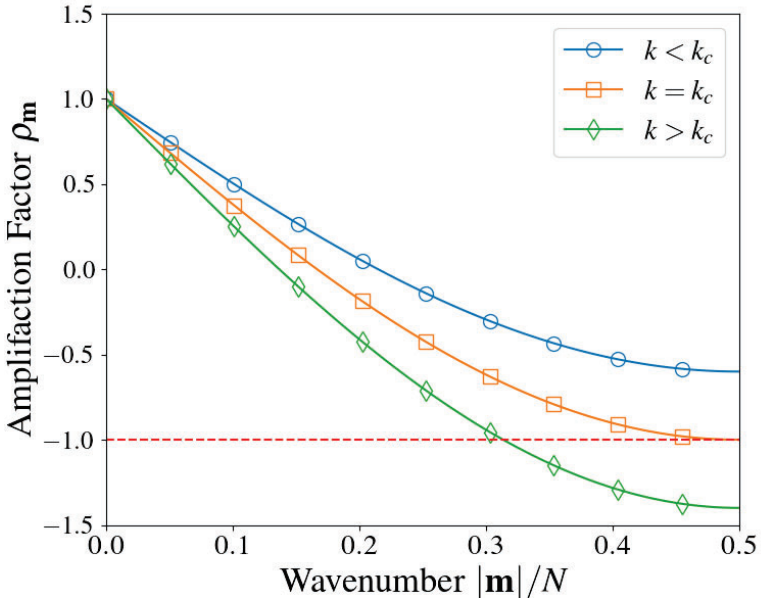


Figure 1. *Forward Euler/central difference scheme (3.12), the amplification factor $\rho_{\mathbf{m}} = 1+kD\lambda_{\mathbf{m}}$ is plotted versus scaled wavenumber $|\mathbf{m}|/N$ for three time step values near the critical value, $k < k_c, k = k_c, k > k_c$; for simplicity the spatial dimension is $d = 1$.*

where the eigenvalues $\lambda_{\mathbf{m}}$ of the discrete Lapacian were given in (3.5). Then the amplification factor of the scheme is obtained as follows,

$$\widehat{u}_{\mathbf{m}}^{n+1} = \frac{1 + \frac{1}{2}kD\lambda_{\mathbf{m}}}{1 - \frac{1}{2}kD\lambda_{\mathbf{m}}}\widehat{u}_{\mathbf{m}}^{n} = \rho_{\mathbf{m}}\widehat{u}_{\mathbf{m}}^{n}, \quad \rho_{\mathbf{m}} = \frac{1 + \frac{1}{2}kD\lambda_{\mathbf{m}}}{1 - \frac{1}{2}kD\lambda_{\mathbf{m}}}. \tag{3.18}$$

Since $\lambda_{\mathbf{m}} \leq 0$, it follows that $|1 + \frac{1}{2}kD\lambda_{\mathbf{m}}| \leq |1 - \frac{1}{2}kD\lambda_{\mathbf{m}}|$, and then $|\rho_{\mathbf{m}}| \leq 1$ for all wavenumbers $\mathbf{m}$ and time steps $k > 0$, so the scheme is unconditionally stable. Figure 2 plots the amplification factor $\rho_{\mathbf{m}}$ of the Crank-Nicholson scheme for the same three time step values as in Figure 1, demonstrating the stability of the scheme for these values. Note also that the Crank-Nicolson scheme (3.16) requires solving a linear system of N^d equations at each time step. However, even though (3.16) is an implicit equation for the numerical solution $u_{\mathbf{j}}^{n+1}$, (3.18) gives an explicit expression for updating the discrete Fourier coefficients which scales like $O(N^d)$, and transforming between the numerical solution and its discrete Fourier coefficients scales like $O(N^d \log N)$ using the Fast Fourier transform (FFT) [1]. Our solution of the linear system utilizes this FFT approach.

3.4. Mass Conservation. To check that both difference schemes satisfy the discrete analogue of mass conservation (2.9), first note that the discrete Laplacian eigenvalue with wavenumber $\mathbf{m} = \mathbf{0}$ is $\lambda_{\mathbf{0}} = 0$. Then the Fourier space equation (3.13) for the Forward Euler scheme

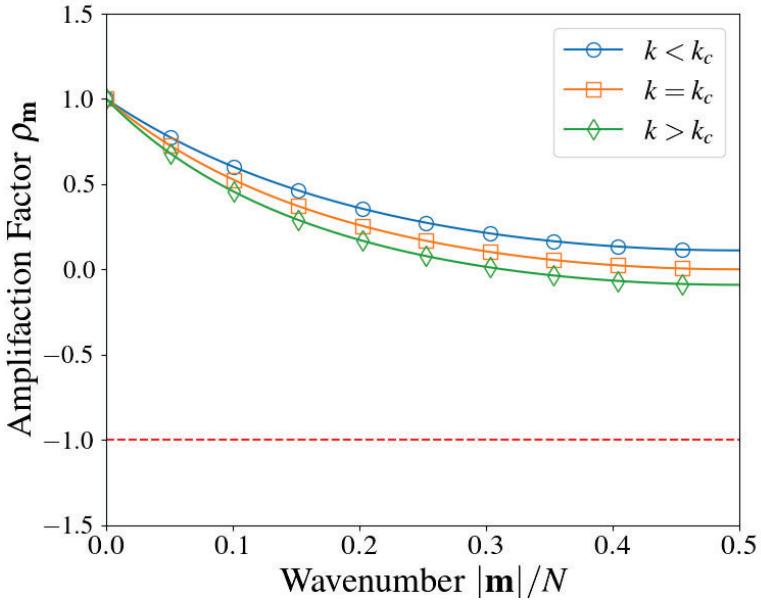


Figure 2. *Crank-Nicolson scheme (3.16), the amplification factor* $\rho_{\mathbf{m}} = (1 + \frac{1}{2}kD\lambda_{\mathbf{m}})/(1 - \frac{1}{2}kD\lambda_{\mathbf{m}})$ *is plotted versus scaled wavenumber* $|\mathbf{m}|/N$ *is plotted for the same time steps* k *as in Figure 1; for simplicity, the spatial dimension is* $d = 1$.

becomes,

$$\widehat{u}_{\mathbf{0}}^{n+1} = \widehat{u}_{\mathbf{0}}^{n} + kD\lambda_{\mathbf{0}}\widehat{u}_{\mathbf{0}}^{n} = \widehat{u}_{\mathbf{0}}^{n},$$

so the average concentration is invariant in time,

$$\frac{1}{N^d}\sum_{\mathbf{j}\in[N]^d} u_{\mathbf{j}}^{n+1} = \frac{1}{N^d}\widehat{u}_{\mathbf{0}}^{n+1} = \frac{1}{N^d}\widehat{u}_{\mathbf{0}}^{n} = \frac{1}{N^d}\sum_{\mathbf{j}\in[N]^d} u_{\mathbf{j}}^{n}. \tag{3.19}$$

Similarly for the Crank-Nicolson scheme (3.16),

$$\widehat{u}_{\mathbf{0}}^{n+1} = \widehat{u}_{\mathbf{0}}^{n} + \frac{kD}{2}(\lambda_{\mathbf{0}}\widehat{u}_{\mathbf{0}}^{n+1} + \lambda_{\mathbf{0}}\widehat{u}_{\mathbf{0}}^{n}) = \widehat{u}_{\mathbf{0}}^{n},$$

which also shows that the average concentration is invariant in time.

4. Molecular Dynamics. The MD simulations presented below utilize the Lennard-Jones potential to simulate the diffusion of argon in helium. These are both noble gases, where argon has atomic number $Z = 18$ and molar mass $m = 39.948$ g/mol, while helium has atomic number $Z = 2$ and molar mass $m = 4.003$ g/mol. This system represents one of the simplest real examples of chemical diffusion [13]. The short-range Lennard-Jones potential is appropriate for modeling the diffusion of noble gases; long-range electrostatic interactions are more expensive to compute, but they can be omitted in the case of small non-polar atoms such

as these. The experimentally determined value of the diffusion coefficient D_{exp} of argon in helium is available [13], which enables verification of the estimation algorithm developed in this work. This section first discusses the equation governing the system evolution, then presents the Lennard-Jones potential, and finally gives the parameters used in the MD simulations.

4.1. Governing equation. The argon and helium gases are represented by a set of N_p particles $\{\mathbf{r}_i \in \mathbb{R}^2, i = 1, \ldots, N_p\}$ moving in two-dimensional space. Newton's equations define the time evolution of the system,

$$m_i \frac{\mathrm{d}^2 \mathbf{r}_i}{\mathrm{d}t} = \mathbf{F}_i, \quad i = 1, \ldots, N_p, \tag{4.1}$$

where m_i is the mass of particle $\mathbf{r}_i$ and $\mathbf{F}_i$ is the force on particle $\mathbf{r}_i$, induced by all the other particles. The total potential energy of the system is the sum of pairwise particle interactions,

$$V_{\text{total}}(\mathbf{r}_1, \ldots, \mathbf{r}_{N_p}) = \sum_{i=1}^{N_p} \sum_{\substack{j=1 \\ j \neq i}}^{N_p} V(r_{ij}), \quad r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|, \tag{4.2}$$

where $V(r)$ is the Lennard-Jones potential given below, r_{ij} is the interparticle distance between $\mathbf{r}_i$ and $\mathbf{r}_j$, and the $j = i$ terms are excluded from the sum because particles do not interact with themselves. The system is considered to be conservative and hence the force on each particle is the gradient of the total potential energy,

$$\mathbf{F}_i = \nabla_{\mathbf{r}_i} V_{\text{total}}(\mathbf{r}_1, \ldots, \mathbf{r}_{N_p}) = \sum_{i=1}^{N_p} \sum_{\substack{j=1 \\ j \neq i}}^{N_p} \nabla_{\mathbf{r}_i} V(r_{ij}), \quad i = 1, \ldots, N_p, \tag{4.3}$$

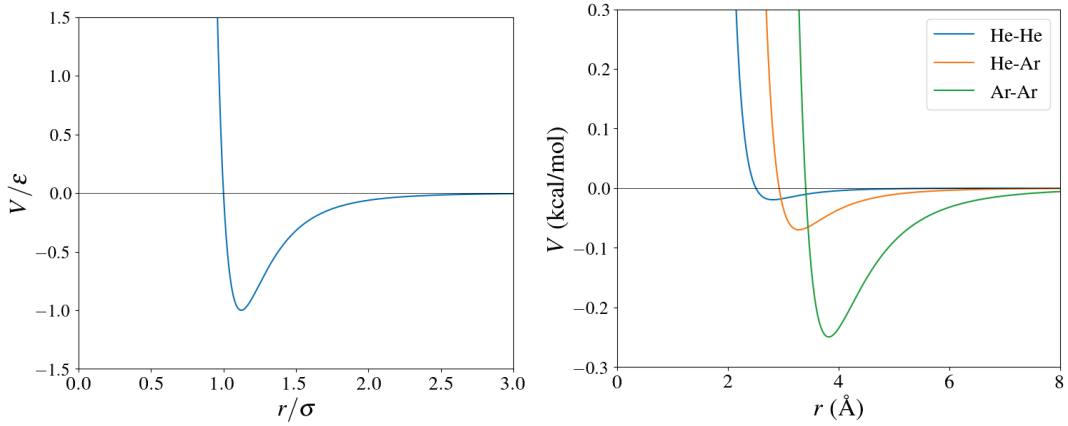
where $\nabla_{\mathbf{r}_i}$ is the gradient with respect to particle $\mathbf{r}_i$.

4.2. Lennard-Jones potential. The Lennard-Jones potential is given by the expression

$$V(r) = \begin{cases} 4\epsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right], & r < r_c, \\ 0, & r \geq r_c, \end{cases} \tag{4.4}$$

where ϵ, σ are parameters, r is the interparticle distance, and r_c is a long-range cutoff. Figure 3(a) plots the scaled Lennard-Jones potential V/ϵ versus the scaled interparticle distance r/σ. The potential characterizes London dispersion forces, where the order 12 term accounts for close-range repulsive forces which dominate when $r < \sigma$, and the order 6 term accounts for long-range attractive forces that operate for $r > \sigma$. The parameter ϵ sets the maximum depth of the potential well which determines the interaction strength, and the parameter σ sets the distance at which the potential $V(r)$ vanishes.

The Lennard-Jones parameters ϵ, σ depend on the physical and chemical properties of the particles in the system. The parameters used in the present work were taken from MD calculations and are displayed in Table 1 [6, 8]. In our simulations, each particle is designated



(a) (b)

Figure 3. *Lennard-Jones potential V given in (4.4), (a) scaled potential V/ϵ versus scaled distance r/σ, (b) dimensional potentials V (kcal/mol) for He-He, Ar-Ar, and mixed Ar-He interactions versus interparticle distance r (Å) with parameters ϵ, σ given in Table 1 [6, 8].*

as either argon or helium, and the interaction potential $V(r)$ in (4.4) between two particles is determined by their identities. There are three types of pairwise interactions: helium/helium, argon/argon, and mixed helium/argon. For each pair of particles, the parameters are selected from Table 1 and Figure 3(b) plots the Lennard-Jones potentials for the three types of pairwise interactions. The relative distance scale of the helium/helium and argon/argon potentials reflects that argon has a larger atomic radius than helium. Similarly, the argon/argon interaction is stronger and has a deeper well, because argon atoms are more polarizable, and exhibit greater induced dipoles.

Interaction	ϵ (kcal/mol)	σ (Å)
He-He	0.0196	2.50
He-Ar	0.0700	2.92
Ar-Ar	0.2498	3.40

Table 1

Lennard-Jones parameters for argon and helium as determined in MD calculations [6, 8], mixed parameter values use the geometric mean of the individual particle parameters. The cutoff distance is $r_c = 20\text{Å}$ for all pairwise interactions.

4.3. LAMMPS. This work utilizes the LAMMPS software package to implement the MD simulations for argon diffusion in helium with the Lennard-Jones potentials described above [9]. The real unit system in LAMMPS is chosen for these simulations, where distance is in Angstroms (Å), energy in kilocalories per mole (kcal/mol), time in femtoseconds (fs = 1e-15 second), temperature in Kelvin (K), and mass in grams per mole (g/mol). The simulations were done in two space dimensions, where the MD simulation box has dimensions 5e4 Å × 5e4 Å, but for simplicity the results are presented using a dimensionless length scale where the unit square $[0, 1]^2$ represents this domain. Periodic boundary conditions were imposed using

the minimum image convention in which the particles are replicated across the sides of the simulation box for the purpose of computing the forces, and upon exiting a side, they re-enter the opposite side. Three MD simulations were performed in which the initial condition in each is a unique pseudo-random seed of 3e4 helium atoms uniformly distributed in the domain $[0,1]^2$ and a patch of 3e4 argon atoms uniformly distributed in the square subdomain $[1/4,3/4]^2$. The initial particle velocities were chosen from the Maxwell-Boltzmann distribution,

$$f(\mathbf{v}) \propto \mathrm{e}^{-\frac{m|\mathbf{v}|^2}{2k_bT}}, \tag{4.5}$$

where $T = 300$ K, m is the mass of a particular atom, and $k_b = 0.001987$ kcal/(mol K) is Boltzmann's constant. The MD timestep was $k_{\mathrm{MD}} = 5$ fs. The order 12 repulsive term of the Lennard-Jones potential (4.4) necessitates a small timestep to ensure sufficient accuracy in the LAMMPS integrator. An NVE integrator was used for the MD simulations, which samples the microcanonical ensemble, where the number of particles, volume, and total energy are fixed. The LAMMPS NVE integrator uses the Störmer-Verlet time stepping scheme.

5. Comparison of MD and FD results. This section compares molecular dynamics and finite-difference simulations of argon diffusion in helium in two space dimensions. Recalling that the MD simulations utilized random initial conditions representing a square patch of argon atoms, the FD calculations were initialized with the following continuum idealization,

$$u_0(\mathbf{x}) = \begin{cases} 1, & \mathbf{x} \in [1/4,3/4]^2, \\ 0, & \mathbf{x} \in [0,1]^2 \setminus [1/4,3/4]^2. \end{cases} \tag{5.1}$$

The total simulation time was 5 ns, where the FD time step was $k_{\mathrm{FD}} = 5$ ps (1e3 time steps) and the MD time step was $k_{\mathrm{MD}} = 5$ fs (1e6 time steps). The MD time step k_{MD} was determined empirically to achieve stable results and is 1000 times smaller than the FD time step k_{FD}; this is because the MD simulations with the explicit Störmer-Verlet time stepping scheme are constrained by the stiffness of the Lennard-Jones potential, whereas the FD calculations use the implicit Crank-Nicolson time stepping scheme which is unconditionally stable. To accommodate the difference between the MD and FD time steps, the fitting procedure discussed in the next section samples the MD trajectory every 1000 time steps.

The simulations were carried out on a single CPU core of an Apple M1 2020 Macbook Pro with 16 GB of memory running the Sonoma 14 operating system. The run time for each FD calculation was less than a second, and the run time for each MD simulation was about 80 minutes. A parallel implementation would reduce the MD run time significantly, and LAMMPS does support parallel computing on multicore systems, but this is beyond the scope of the present work.

Figure 4 compares the atomistic MD results and the continuum FD results computed with grid spacing $h = 1/50$. The diffusion coefficient in the FD simulations ($D = 0.7948$ cm^2/s) is the optimal value determined by the estimation algorithm described in the next section. The time is given at the top of each column. Row 1 plots the MD results, where the argon atoms are red dots and the helium atoms are a uniform blue background. As time proceeds, the initial square distribution of argon atoms evolves into an expanding radially symmetric distribution.

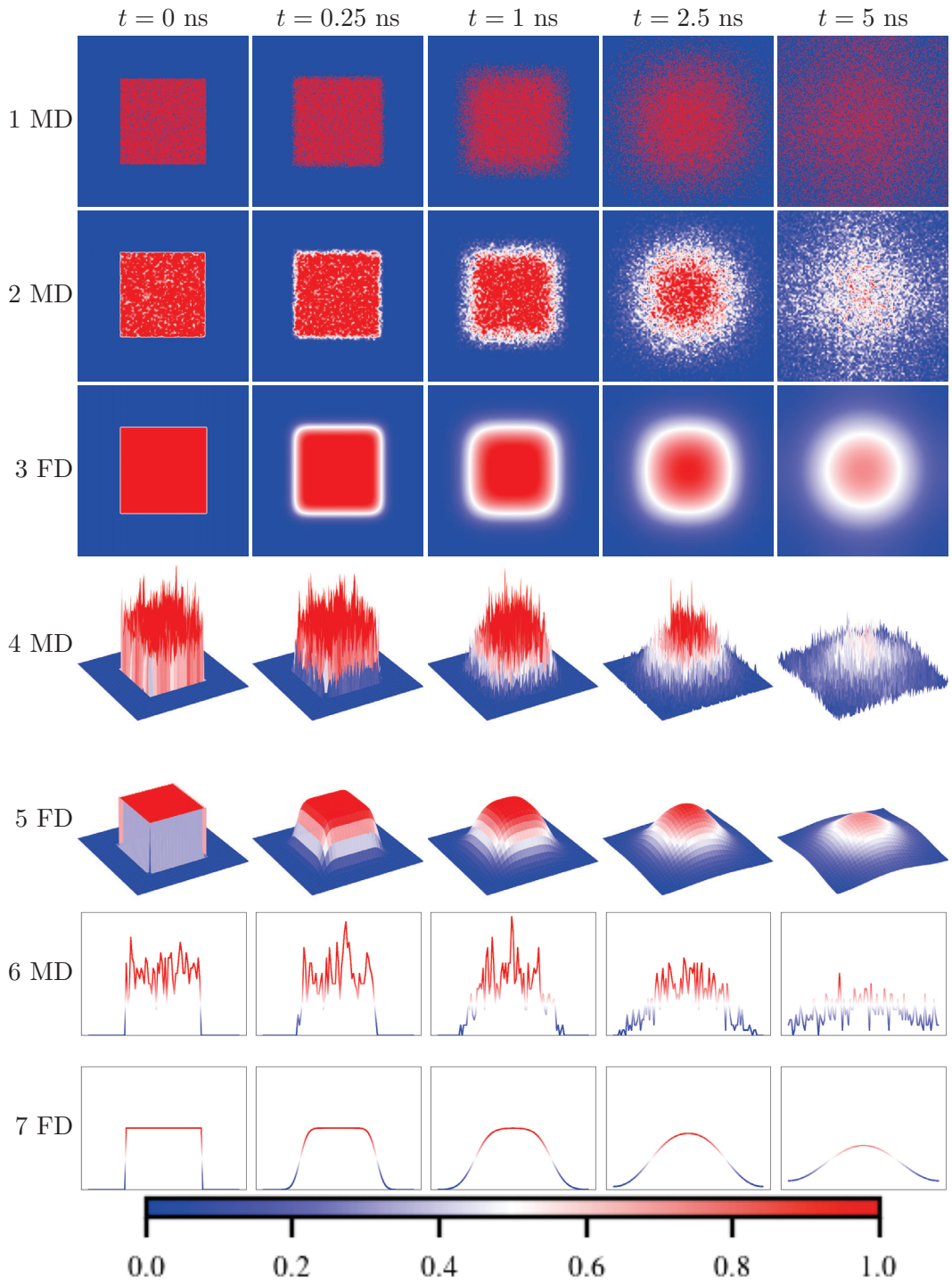


Figure 4. *Numerical results for diffusion of argon in helium, color bar indicates normalized argon concentration, MD binning parameter $N = 50$, FD solution of diffusion equation (2.4) computed on a 50×50 grid, Columns 1-5 display times $t = 0, 0.25, 1, 2.5, 5$ ns, Row 1: MD snapshots of argon atoms, Rows 2, 4, 6: binned MD concentration, Rows 3, 5, 7: computed FD concentration.*

To facilitate comparison of the MD and FD results, the MD argon trajectories $\{\mathbf{r}_i^n\}$ were converted into a concentration $U_{\mathbf{j}}^n$ defined on the FD grid by a binning process [3] that counts the number of argon atoms in each FD grid cell, $N_{\mathbf{j}}^n$, and scales the count to lie between 0 and 1,

$$N_{\mathbf{j}}^n = \#\{\mathbf{r}_i^n \in \text{grid cell } \mathbf{j}\}, \quad U_{\mathbf{j}}^n = \frac{N_{\mathbf{j}}^n}{\max N_{\mathbf{j}}^n}. \tag{5.2}$$

Row 2 in Figure 4 plots the resulting binned MD argon concentration $U_{\mathbf{j}}^n$ and Row 3 plots the corresponding FD concentration $u_{\mathbf{j}}^n$ as heat maps on the FD grid. Although the MD results are noisy and the FD results are smooth, there is a clear similarity between the two concentrations. The similarity is also seen in Rows 4 and 5 which present surface plots of the MD and FD concentrations. Rows 6 and 7 show the $x = 1/2$ cross-section of the MD and FD concentrations. In the surface plots and cross-sections, the MD and FD maximum argon concentrations decrease at similar rates. Figure 4 shows that using an optimal value of the diffusion coefficient D in the FD calculation ensures a close connection between the FD and MD results.

One may consider how the match between the MD and FD results depends on the binning parameter N, which also determines the FD grid size, $h = 1/N$. The accuracy of the FD solution $u_{\mathbf{j}}^n$ relative to the true solution of the diffusion equation $u(\mathbf{x}, t)$ increases with N. Additionally, the binned MD concentration $U_{\mathbf{j}}^n$ will better resolve the computed MD trajectories $\mathbf{r}_i^n$ for large N. However as the binning parameter N increases, the binned MD concentration will become increasingly noisy due to the use of a finite number particles in the MD simulation. Based on these competing effects, we expect there to be an optimal level of binning resolution to achieve the best match between the MD and FD results. The balance between increasing resolution and noise also depends on several other factors including the MD domain size, number of MD particles, and initial MD distribution.

6. Diffusion coefficient estimation algorithm. This section describes the algorithm for estimating the optimal diffusion coefficient D_{opt} for the continuum diffusion equation (2.4). This is essentially a least-squares fitting procedure that extracts D_{opt} from the MD trajectories. Recall that the MD trajectories $\mathbf{r}_i^n$ give the position of each argon and helium atom at a sequence of discrete time steps, and these trajectories are represented on the FD grid by the binned MD argon concentration $U_{\mathbf{j}}^n$ obtained using the previously described binning process. Let $u_{\mathbf{j}}^n(D)$ be the numerical solution of the continuum diffusion equation (2.4) with diffusion coefficient D given by the Crank-Nicolson scheme (3.16). The vector of residuals is given by

$$\{U_{\mathbf{j}}^n - u_{\mathbf{j}}^n(D)\}, \quad n = 0, \ldots, n_{\max}, \quad \mathbf{j} \in [N]^2, \tag{6.1}$$

and the cost function is the mean square deviation between the MD and FD results,

$$\texttt{cost}(D) = \frac{1}{N^2} \sum_{n=0}^{n_{\max}} \sum_{\mathbf{j} \in [N]^2} (U_{\mathbf{j}}^n - u_{\mathbf{j}}^n(D))^2. \tag{6.2}$$

Minimizing the cost function is a nonlinear optimization problem that yields the optimal diffusion coefficient, $D_{\text{opt}}(N)$. Figure 5 plots the cost function for the argon-helium system

with MD binning parameter $N = 20, 50, 100$, showing that $\texttt{cost}(D)$ is a smooth convex function with a unique global minimum $D_{\text{opt}}(N)$ in each case.

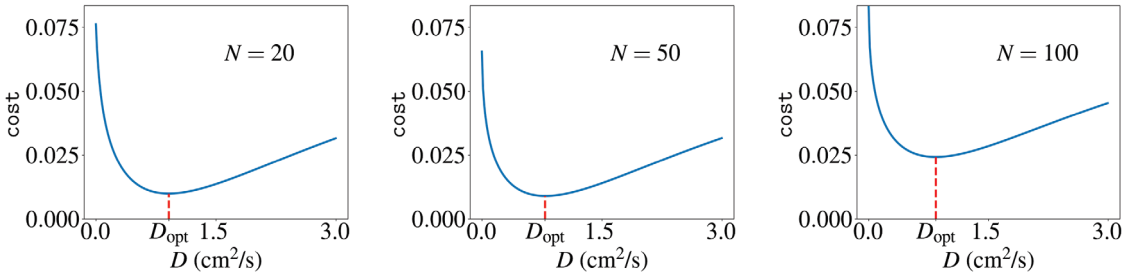


Figure 5. *The cost function* $\texttt{cost}(D)$ (6.2) *of the diffusion coefficient estimation algorithm is plotted versus the diffusion coefficient* D *near the global minimum* $D_{opt}(N)$ *with MD binning parameters* $N = 20, 50, 100$.

Algorithm 6.1 shows the steps for computing the optimal diffusion coefficient,

$$D_{\text{opt}}(N) = \arg\min_D \texttt{cost}(D) = \arg\min_D \frac{1}{N^2} \sum_{n=0}^{n_{\max}} \sum_{\mathbf{j}\in[N]^d} (U_{\mathbf{j}}^n - u_{\mathbf{j}}^n(D))^2. \tag{6.3}$$

Lines 1-3 input the MD argon trajectories $\mathbf{r}_i^n$, MD binning parameter N, and an initial guess for the diffusion coefficient D_0. Line 4 computes the binned MD concentration, $U_{\mathbf{j}}^n$, using the binning procedure described in (5.2). Line 5 computes the optimal diffusion coefficient $D_{\text{opt}}(N)$ using the `curve_fit` routine from the Python package `scipy` [10] to solve the least-squares minimization problem, Line 6 returns $D_{\text{opt}}(N)$. Next we explain how the `curve_fit` routine works.

Algorithm 6.1 Diffusion Coefficient Estimation Algorithm

1: **input** MD trajectories, $\mathbf{r}_i^n, i = 1, \ldots, N_p$, time steps $n = 0, \ldots, n_{\max}$
2: **input** MD binning parameter, N
3: **input** initial diffusion coefficient guess, D_0
4: compute binned MD concentration, $U_{\mathbf{j}}^n$, from (5.2)
5: compute $D_{\text{opt}}(N)$ from (6.3) using the `curve_fit` routine from `scipy`
6: **return** $D_{\text{opt}}(N)$

The `curve_fit` routine uses the Levenberg-Marquardt (LM) algorithm [4] to compute the optimal diffusion coefficient $D_{\text{opt}}(N)$. The input for `curve_fit` is an initial guess D_0 for the diffusion coefficient. The LM algorithm is an iterative scheme that combines the stability of gradient descent for minimizing the cost function $\texttt{cost}(D)$ with the accuracy of Newton's method for finding a critical point of $\texttt{cost}(D)$. At each step, the diffusion coefficient is updated, $D \to D + \delta D$, where the increment δD is the solution of the normal equations,

$$(\mathbf{J}^{\mathrm{T}}\mathbf{J} + \lambda)\delta D = \mathbf{J}^{\mathrm{T}}(U_{\mathbf{j}}^n - u_{\mathbf{j}}^n(D)), \tag{6.4}$$

that arise from linearizing the least-squares minimization problem (6.3). In these equations, $\mathbf{J} = \partial u_{\mathbf{j}}^n / \partial D$ is the Jacobian of the residual vector, which is computed numerically, and λ is

an adjustable regularization parameter. When the current step causes a large reduction in $\texttt{cost}(D)$, λ is set to a small value and the algorithm is close to Newton's method, but if the reduction in $\texttt{cost}(D)$ is small, then λ is increased to bring the iteration closer to gradient descent. The `curve_fit` routine utilizes a procedure `resid` which takes a diffusion coefficient D as input, runs the Crank-Nicolson FD scheme (3.16) with this D, and returns the residual vector $\{U_{\mathbf{j}}^n - u_{\mathbf{j}}^n(D)\}$. Note that even though the explicit forward Euler FD scheme (3.12) is faster, the implicit Crank-Nicolson scheme (3.16) was chosen to avoid any stability issues that might arise in exploring the diffusion coefficient parameter space.

In general, for a single-parameter least-squares optimization problem with a global minimum as in the present case, the LM algorithm is known to converge to a local minimum, assuming the initial guess is sufficiently close to the minimum [4]. In the present case, we tested various initial guesses and convergence was obtained in all cases, probably owing to the smooth convex nature of the cost function shown in Figure 5.

7. Diffusion coefficient estimates. The procedure outlined in Section 6 was used to compute estimates $D_{\text{opt}}(N)$ for the diffusion coefficient of argon in helium for several binning parameter values $N = 20, 50, 100$. The MD simulations were done with three different initial random seeds and the estimates of $D_{\text{opt}}(N)$ were averaged. The FD calculations used grid spacing $h = 1/N$. Table 2 shows the results, where column 1 gives the MD binning parameter N, column 2 gives the estimated $D_{\text{opt}}(N)$ in cm^2/s, column 3 gives the value of the cost function at $D_{\text{opt}}(N)$ as defined in (6.2), column 4 gives the 95% confidence interval for $D_{\text{opt}}(N)$ as reported by `curve_fit`. The first three rows give the results of the estimation procedure for the different MD binning parameters, and the last row gives the experimentally measured diffusion coefficient which was obtained using gas chromatography [13].

MD binning parameter	$D_{\text{opt}}(N)$ [cm^2/s]	$\texttt{cost}(D_{\text{opt}}(N))$	95% CI
$N = 20$	0.9121	0.009870	1.48e-3
$N = 50$	0.7948	0.008930	0.54e-3
$N = 100$	0.8417	0.024214	0.46e-3
experiment [13]	0.7335		8.70e-3

Table 2

Results of diffusion coefficient estimation for argon in helium, column 1: MD binning parameter N, column 2: optimal diffusion coefficient $D_{\text{opt}}(N)$ averaged over three MD runs with different initial random seeds, column 3: cost function $\texttt{cost}(D_{\text{opt}}(N))$ as defined in (6.2), column 4: 95% confidence interval, last row gives experimental value D_{exp} of the diffusion coefficient [13].

The following observations were made.

- Column 2 shows that the estimated diffusion coefficient $D_{\text{opt}}(N)$ agrees best with experiment for the intermediate MD binning parameter $N = 50$. This can be understood as follows. The smaller MD binning parameter $N = 20$ makes the FD grid spacing $h = 1/N$ larger, and then the FD result $u_{\mathbf{j}}^n$ is less accurate. On the other hand, the larger MD binning parameter $N = 100$ makes the bins smaller and hence they contain fewer atoms, and then the MD binning concentration $U_{\mathbf{j}}^n$ has more statistical noise.
- Column 3 shows that the cost function achieves its smallest value for $N = 50$, which is consistent with the previous point that the best agreement between model and

experiment occurs for $N = 50$.

- Column 4 shows that the model confidence interval becomes narrower with higher resolution; however, the experimental confidence interval is much wider, which perhaps can be attributed to experimental measurement uncertainty.

As shown in Table 2, although the estimated diffusion coefficient $D_{\text{opt}}(50)$ is close to the experimental value, there is still a discrepancy which may be attributed to several factors. One factor is the choice of numerical parameters including the number of atoms in the MD simulation N_p, binning parameter N, and time steps $k_{\text{MD}}, k_{\text{FD}}$ in the MD and FD calculations. A second factor is that the MD and FD calculations were initialized differently and this could affect the quality of the estimation results. A third factor is that the Lennard-Jones potential utilized in the MD simulations is only an approximation to the true interparticle interaction between the gas molecules. A fourth factor is that the calculations were carried out in two-dimensional space, while there may be three-dimensional effects in the experiment. A final factor is the statistical noise in the binned MD data which could be reduced with a convolution filter [3]. These are all possible directions for future investigation.

Finally, we emphasize our choice to use a small sample of only three MD simulations at each N value, considering the relatively high MD runtime (up to 80 minutes). We anticipate the utility of our method to be the extraction of an accurate diffusion coefficient estimate from a relatively small amount of MD data, which may be applied to a efficiently simulate a larger continuum system. Using more simulations may reduce the impact of the statistical noise observed in the binned MD concentration of Figure 4. However, we believe the reasonable agreement of the estimated and experimental diffusion coefficients justifies the method.

8. Summary. This work studied the diffusion of a patch of argon gas in a background of helium gas with the goal of determining the diffusion coefficient of the argon atoms. Two models were utilized, (1) finite difference (FD) simulations of the continuum diffusion equation (2.4) with a given diffusion coefficient D, and (2) molecular dynamics (MD) simulations where the gas atoms were represented by point particles interacting through the Lennard-Jones potential (4.4). The MD simulations were taken as the true physical representation of the diffusion process. To estimate the argon diffusion coefficient, the MD trajectories of the argon particles were converted into a binned concentration on the FD grid, and the optimal diffusion coefficient D_{opt} was obtained by minimizing the least squares difference between the FD concentration and the binned MD concentration. Comparison of the numerical results with the experimental diffusion coefficient suggests that our optimization estimation approach was successful, although there is still a discrepancy and several directions for future investigation were suggested.

The present work explored a system where an argon patch diffused into a background of helium. Another interesting physical system is a mixture of polymers, where both phase-separation and diffusion dynamics are present. Such systems may be modeled by the Allen-Cahn and Cahn-Hilliard equations,

$$\frac{\partial u}{\partial t} = M[\gamma\nabla^2 u - u^3 + u], \quad \frac{\partial u}{\partial t} = M\nabla^2[u^3 - u - \gamma\nabla^2 u], \tag{8.1}$$

where u is the order parameter giving the purity of phases, M is the mobility coefficient, and

$\sqrt{\gamma}$ sets the width of the transition zone between the phases at equilibrium. The parameter estimation algorithm developed in this article may also be applied to find optimal values of M and γ. Some preliminary results for the phase separation equations are contained in the REU project report [11].

Acknowledgments. The author thanks the University of Michigan Department of Mathematics for hosting this project which was supported by National Science Foundation grant DMS-2110767.

REFERENCES

[1] M. I. M. Copetti and C. M. Elliott, *Kinetics of Phase Decomposition Processes: Numerical Solutions to Cahn–Hilliard Equation*, Mater. Sci. Technol., 6 (1990), pp. 273–284, https://doi.org/10.1179/mst.1990.6.3.273.

[2] A. Fick, *Ueber Diffusion*, Ann. Phys., 170 (1855), pp. 59–86, https://doi.org/10.1002/andp.18551700105.

[3] R. G. Larson, T. T. Perkins, D. E. Smith, and S. Chu, *Hydrodynamics of a DNA molecule in a flow field*, Phys. Rev. E, 55 (1997), pp. 1794–1797, https://doi.org/10.1103/PhysRevE.55.1794, https://link.aps.org/doi/10.1103/PhysRevE.55.1794.

[4] D. W. Marquardt, *An Algorithm for Least-Squares Estimation of Nonlinear Parameters*, Journal of the Society for Industrial and Applied Mathematics, 11 (1963), pp. 431–441, http://www.jstor.org/stable/2098941 (accessed 2025-02-05).

[5] F. Melzer, R. Breuer, R. Dahlmann, and C. Hopmann, *Calculating diffusion coefficients from molecular dynamics simulations for foam extrusion modelling of polypropylene with CO2, N2 and ethanol*, J. Cell. Plast., 58 (2022), pp. 603–622, https://doi.org/10.1177/0021955X221087598, https://doi.org/10.1177/0021955X221087598, https://arxiv.org/abs/https://doi.org/10.1177/0021955X221087598.

[6] R. A. Messerly, N. Gokul, A. J. Schultz, D. A. Kofke, and A. H. Harvey, *Molecular calculation of the critical parameters of classical helium*, J. Chem. Eng. Data, 65 (2019).

[7] P. Olver, *Introduction to Partial Differential Equations*, Undergraduate Texts in Mathematics, Springer International Publishing, 2013, https://books.google.com/books?id=aQ8JAgAAQBAJ.

[8] A. Rahman, *Correlations in the Motion of Atoms in Liquid Argon*, Phys. Rev., 136 (1964), pp. A405–A411, https://doi.org/10.1103/PhysRev.136.A405, https://link.aps.org/doi/10.1103/PhysRev.136.A405.

[9] A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. in 't Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen, R. Shan, M. J. Stevens, J. Tranchida, C. Trott, and S. J. Plimpton, *LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales*, Comp. Phys. Comm., 271 (2022), p. 108171, https://doi.org/10.1016/j.cpc.2021.108171.

[10] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, *SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python*, Nat. Methods, 17 (2020), pp. 261–272, https://doi.org/10.1038/s41592-019-0686-2.

[11] I. G. Viviano, *Comparison of continuum and atomistic models for chemical diffusion and phase separation*, (2024), https://lsa.umich.edu/content/dam/math-assets/reu-su22/reu-2024/viviano_REU_2024.pdf.

[12] S. Wasik and K. McCulloh, *Measurements of Gaseous Diffusion Coefficients by a Gas Chromatographic Technique*, J. Res. Natl. Bur. Stand., 73A (1969), p. 207, https://doi.org/10.6028/jres.073a.018.

[13] S. P. Wasik and K. E. McCulloh, *Measurements of Gaseous Diffusion Coefficients by a Gas Chromatographic Technique*, J. Res. Natl. Bur. Stand. A Phys. Chem., 73A (1969), pp. 207–211.