Bayesian Inference of Atomic Diffusivity in a Binary Ni/Al System Based on Molecular Dynamics

MULTISCALE MODEL. SIMUL.
Vol. 9, No. 1, pp. 486–512

BAYESIAN INFERENCE OF ATOMIC DIFFUSIVITY IN A BINARY NI/AL SYSTEM BASED ON MOLECULAR DYNAMICS*

F. RIZZI†, M. SALLOUM†, Y.M. MARZOUK‡, R.-G. XU§, M. L. FALK¶, T. P. WEIHS¶, G. FRITZ¶, AND O. M. KNIO†

Abstract. This work focuses on characterizing the integral features of atomic diffusion in Ni/Al nanolaminates based on molecular dynamics (MD) computations. Attention is focused on the simplified problem of extracting the diffusivity, D, in an isothermal system at high temperature. To this end, a mixing measure theory is developed that relies on analyzing the moments of the cumulative distribution functions (CDFs) of the constituents. The mixing measures obtained from replica simulations are exploited in a Bayesian inference framework, based on contrasting these measures with corresponding moments of a dimensionless concentration evolving according to a Fickian process. The noise inherent in the MD simulations is described as a Gaussian process, and this hypothesis is verified both a priori and using a posterior predictive check. Computed values of D for an initially unmixed system rapidly heated to 1500 K are found to be consistent with experimental correlation for diffusion of Ni into molten Al. On the contrary, large discrepancies with experimental predictions are observed when D is estimated based on large-time mean-square displacement (MSD) analysis, and when it is evaluated using the Arrhenius correlation calibrated against experimental measurements of self-propagating front velocities. Implications are finally drawn regarding extension of the present work and potential refinement of continuum modeling approaches.

Key words. Reactive multilayers, atomic diffusion, molecular dynamics, Bayesian interference

AMS subject classifications. 74A65, 62F15, 62C10, 35K05, 65C40, 74A25

DOI. 10.1137/10080590X

1. Introduction.

Reactive multilayer foils are nanostructures comprised of alternating layers of materials that mix exothermically. The microstructure of these foils is characterized in terms of the bilayer thickness, defined as the thickness of a pair of alternating material layers. While a typical bilayer thickness is small, on the order of tens of nanometers, a foil can contain a large number of layers, yielding a total thickness ranging from 10 to 200 μm. The reaction in these multilayer foils can be triggered through a concentrated energy pulse, such as an electric spark or a

*Received by the editors August 20, 2010; accepted for publication (in revised form) January 7, 2011; published electronically March 31, 2011.
http://www.siam.org/journals/mms/9-1/80590.html

†Department of Mechanical Engineering, Johns Hopkins University, Baltimore, MD 21218 (frizzi1@jhu.edu, mnsallo@sandia.gov, knio@jhu.edu). The last author was supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Materials Sciences and Engineering under award DE-SC0002509. The first author was supported by the U.S. Department of Energy, Office of Advanced Scientific Computing Research under award DE-SC0002506. The second author was supported by Office of Naval Research under award N00014-07-1-0740 (GF).

‡Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA 02139 (ymarz@mit.edu). This author was supported by the U.S. Department of Energy, Office of Advanced Scientific Computing Research and Sandia National Laboratories under contract agreement 971321.

§Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218 (XRg117@pha.jhu.edu). This author was supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Materials Sciences and Engineering under award DE-SC0002509.

¶Department of Materials Science and Engineering, Johns Hopkins University, Baltimore, MD 21218 (mfalk@jhu.edu, weihs@jhu.edu, grey.fritz@jhu.edu). The first and second author were supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Materials Sciences and Engineering under award DE-SC0002509. The last author was supported by Office of Naval Research under award N00014-07-1-0740 (GF).

thermal pulse. Self-propagating fronts consequently occur, with speeds ranging from 1 to 30 m/s [1, 2, 3, 4, 5, 6].

Several experimental and computational studies have aimed at studying the mean velocity of self-propagating fronts in reactive multilayers [7, 8, 9, 10, 11, 12, 13, 14]. In particular, models for self-propagating fronts have been based on a simplified, continuum description of atomic mixing, assuming a Fickian diffusion process with a diffusivity exhibiting an Arrhenius dependence on the temperature, T. Thus, the diffusivity is typically expressed as

$$D(T) = D_0 \exp\left(-\frac{E}{RT}\right), \tag{1}$$

where D_0 is the pre-exponent, E is the activation energy, and R is the gas constant. In this formalism, D_0 and E are regarded as empirical constants, which are obtained from best fits to experimental observations of front velocities [7].

The evolution of the fronts, mainly steady, can also show unsteady features [10, 11, 15, 16, 17]. Existing models of self-propagating fronts in reactive multilayers have been based on steady formulations of the multidimensional reaction fronts, thus neglecting all the unsteady effects. This assumption limits the computational cost required to obtain the desired predictions.

The most common approach used in continuum modeling of reactive multilayer is based on single-layer, adiabatic formulations [7, 9, 13, 18], in which computations are typically restricted to a single representative bilayer of the multilayer stack. These models assume that atomic mixing can be described in terms of a Fickian process, and that a fast, diffusion-limited combustion regime prevails. Moreover, since the thermal diffusivity is orders of magnitude larger than the atomic diffusivity, temperature variations across the layers are ignored so that the temperature distribution varies only with axial position and time. To further reduce the computational cost, symmetry arguments have been successfully applied to concentrate the computations to a quarter bilayer. The assumption of an adiabatic system has been relaxed in some studies, which have focused on the evolution of self-propagating fronts in thin foil geometries. In this case, the curvature of the reaction front is neglected and the heat losses can be accounted for without the need to develop a multidimensional formulation. The resulting lumped parameter models have been particularly used to explore the impact of conductive and radiative losses on the front velocity [12], and to characterize the evolution of the front in reactive foils deposited on a substrate [19].

In recent work [20, 21, 22], Salloum and Knio developed and tested a reduced model for simulating transient reactions in Ni/Al nanolaminates, based on the continuum formulation developed in [13]. The approach introduced in [20] aims at reducing the computational costs involved in transient computations, namely by expressing the local atomic mixing rates in terms of an evolution equation for a stretched time scale that reflects the age of the mixed layer. This is obtained based on an approximate theoretical analysis of a quasi-one–dimensional (1D) evolution equation of the conserved scalar. The analysis yields a canonical solution for the concentration profile, whose instantaneous moments can be readily determined based on the evolution of the stretched time scale. As shown in [20, 21, 22], this leads to substantial computational savings, and can thus be exploited for simulations of transient reaction fronts in multiple space dimensions.

The continuum approaches described above have been successful in capturing a variety of experimentally observed phenomena, including the dependence of the front velocity and ignition properties on material and microstructural parameters.

A weakness or criticism of prior approaches, however, concerns the fact that atomic mixing rates are based on empirical fits, which are calibrated based on a limited set of experimental data and reflect averages over wide temperature ranges. This leads to fundamental questions concerning the range of validity of the corresponding computational predictions. One of the goals of the present work is to explore the potential for molecular dynamics (MD) computations to shed additional light on the integral features of atomic mixing processes, and consequently yield improved or more robust continuum representations for these processes. Another motivation concerns the possibility of using MD models to explore material systems that have not been experimentally observed, namely before substantial investments are applied towards synthesizing or fabricating initial samples. A long-term perspective is exploiting MD computations in a new multiscale paradigm that spans multiple decades of spatial length-scales, namely by extracting key parameters of reduced-order models directly and thus bypassing intermediate continuum or mesoscale regimes.

The objective of the present work is to characterize diffusion in an Ni/Al system using MD computations in conjunction with a Bayesian inference formalism. Being our first attempt in this direction, attention is focused on the simplified problem of extracting the diffusivity in an Ni/Al system under isothermal-isobaric conditions. This simplified approach allows us to look at the extraction of a single, rather than temperature-dependent diffusivity and, consequently, test the suitability of the inference formalism in an idealized setting.

Previous experiments have aimed at studying the diffusivity of Ni into molten Al under different conditions; see, e.g., [23, 24, 25, 26] and references therein. In particular, Zhao et al. [23] investigated the dissolution kinetics of nickel in liquid aluminum under the influence of a DC-current. The current has a marked effect on the diffusion constant and, correspondingly, the application of the current significantly decreased the activation energy of dissolution. In the context of numerical studies [27, 28], the Voter and Chen version of the embedded atom model (VC-EAM) [29] has been used in conjunction with MD to compute the diffusion constant D and shear viscosity of the FCC transition metals Ni, Pd, Pt, Cu, Ag, and Au in the liquid phase. The VC-EAM differs from other EAM approaches in that the core-core pair interaction has a medium-range attractive contribution, rather than being entirely repulsive, and the properties of the diatomic molecule, as well as bulk solid-state properties, are used in fitting the embedding function and pair interaction. The good agreement between the VC-EAM predictions and available experimental data [30] suggests that this model and other EAM approaches may also be capable of describing reasonably well the transport coefficients of binary liquid alloys including at least one transition metal. Comparison between simulated and experimental data for the transport coefficients can be made in the case of liquid Ni-Al, for which experimental data is available; see [31] and references therein. The above studies, however, make use of common, well-established methods to extract the diffusivity, e.g., the Einstein–Smoluchowsky diffusion equation, in which the diffusion coefficient appears in the slope of the mean-square displacement versus time, or the Green–Kubo relation, which expresses the transport coefficient in terms of the velocity autocorrelation function.

In the present study, we are primarily interested in describing the evolution of the system from a state consisting of compositionally pure layers towards a well-mixed state, and, therefore, an alternative approach is used. Specifically, we develop a mixing measure theory for the atomistic system, based on analyzing the moments of the cumulative distribution function (CDF) of the atoms. The mixing measure thus

obtained is compared to corresponding moments in the continuum Fickian formulation of mixing, following the analysis in [20]. Replica MD computations are used to generate a set of data to which a Bayesian inference analysis is applied to extract the diffusivity, D. This formalism allows us to rationally account for the noise inherent in the MD computations and, consequently, establish well-defined posterior estimates of the diffusion coefficient.

The article is organized as follows. In section 2, we provide a brief summary of the continuum formulation introduced in [20], and, then, derive analytical expressions for the appropriate moments of a dimensionless concentration. In section 3, we outline the MD computations and introduce the mixing measure associated with the atomistic system. In section 4, we describe the Bayesian inference approach used to extract the diffusivity. Results and conclusions are presented in sections 5 and 6, respectively.

2. Continuum model. In this section, we discuss the continuum model adopted to describe mixing. This approach is based on the recent work by Salloum and Knio [20], where a reduced-order model was developed to simulate transient reactions in Ni/Al nanolaminates.

Figure 1 shows a schematic of a representative bilayer of Ni/Al of total thickness λ. Assuming that the diffusion process is Fickian and that diffusion across the layers is dominant (x-direction in Figure 1), atomic mixing is described in terms of a 1D evolution equation of a dimensionless conserved scalar, $c(x,t)$, according to

$$\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2}, \tag{2}$$

where D is the diffusion coefficient. The conserved scalar is defined so that $c = 1$ for pure Al, $c = -1$ for pure Ni, and $c = 0$ for pure NiAl. The layers are assumed to be geometrically flat and organized in a periodic arrangement. We also assume that the layers are initially compositionally pure, i.e., composed of either Al or Ni.

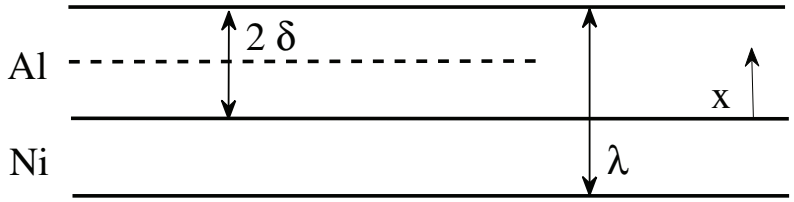


FIG. 1. *Schematic illustration of a representative Al/Ni bilayer, where λ is the bilayer thickness, 2δ is thickness of the Aluminum layer, and x is the spatial coordinate.*

To simplify the analysis [20], the domain of interest is restricted to the region $0 \le x \le \delta$, where δ is the half-thickness of the Al layer, as shown in Figure 1. At $x = 0$, i.e., at the Ni/Al interface, a state of complete mixing prevails, and so the Dirichlet condition $c = 0$ is imposed. Meanwhile, the approximate symmetry condition $\partial c/\partial x = 0$ is imposed at $x = \delta$, i.e., at the upper bound of the domain.

Following the analysis by Salloum and Knio [20], a canonical solution of (2) is sought by introducing the normalized stretched time variable,

$$\tau \equiv \int_0^t \frac{D(T)}{\delta^2} dt, \tag{3}$$

and the normalized spatial coordinate,

$$\xi \equiv x/\delta. \tag{4}$$

Note that for the present isothermal computations, the expression for the stretched time variable simplifies to

$$\tau = Dt/\delta^2. \tag{5}$$

The definitions above enable us to rewrite (2) as

$$\frac{\partial c(\xi,\tau)}{\partial \tau} = \frac{\partial^2 c(\xi,\tau)}{\partial \xi^2} \tag{6}$$

with initial condition

$$c(\xi,0) = g(\xi) = 1 \qquad \forall\ \xi \in [0,1], \tag{7}$$

and boundary conditions

$$c(0,\tau) = 0 = \frac{\partial c}{\partial \xi}(1,\tau) \qquad \forall\ \tau > 0. \tag{8}$$

The advantage of the transformation above is that both the diffusion coefficient, D, and the domain size, δ, drop out from the formulation of the canonical system. Hence, a single solution can be obtained, and later generalized to arbitrary values of D and δ.

The solution of the canonical problem can be obtained analytically in terms of a Fourier-sine series

$$c(\xi,\tau) = \sum_{n=1}^{\infty} F_n \exp\left[-\left(\frac{(2n-1)\pi}{2}\right)^2 \tau\right] \sin\left(\frac{(2n-1)\pi}{2}\ \xi\right), \tag{9}$$

where

$$F_n = \frac{4}{(2n-1)\pi} \tag{10}$$

are the Fourier coefficients. This solution can be readily exploited to estimate the first and second moments of the concentration profile, which can be respectively expressed as

$$\mathcal{M}_c(t) \equiv \frac{1}{\delta}\int_0^{\delta} c(x,t)\ dx = \sum_{n=1}^{\infty} \exp\left[-\left(\frac{(2n-1)\pi}{2}\right)^2 \frac{D}{\delta^2}\ t\right] \frac{8}{(2n-1)^2\pi^2}, \tag{11}$$

and

$$\overline{c^2}(t) \equiv \frac{1}{\delta}\int_0^{\delta} c^2(x,t)\ dx = \sum_{n=1}^{\infty} \frac{(F_n)^2}{2} \exp\left[-\frac{(2n-1)^2\pi^2}{2}\frac{D}{\delta^2}\ t\right], \tag{12}$$

where we have substituted the simplified expression of τ from (5) in order to make explicit the dependence of the moments on the dimensional variables t, δ, and D.

Note that the first moment $\mathcal{M}_c$ defines a measure that can be used to quantify the amount of mixing. Specifically, for the unmixed initial state we have

$$c(x,t=0) = 1 \qquad \Longrightarrow \qquad \mathcal{M}_c(t=0) = 1, \tag{13}$$

whereas a completely mixed system yields

$$c(x, t \to \infty) = 0 \qquad \Longrightarrow \qquad \mathcal{M}_c(t \to \infty) = 0. \tag{14}$$

Note, also, that $\mathcal{M}_c(t)$ decays monotonically from $\mathcal{M}_c = 1$ towards its asymptotic value, $\mathcal{M}_c = 0$, at a rate that depends on the diffusion constant D.

In the analysis below, we will be primarily interested in characterizing the integral features of the mixing process, namely for intermediate and large values of the dimensionless timescales, i.e., $\tau = \emptyset(1)$ or larger. In other words, we are interested in time scales over which the reactants are substantially consumed. In these situations, the Fourier-sine series given above is quite efficient, in the sense that accurate solutions can be obtained when the series is truncated after the first few terms. In fact, as the dimensionless time, τ, becomes large, the Fourier-sine series asymptotes to the corresponding leading term yielding

$$c(\xi, \tau) \sim \frac{4}{\pi} \exp\left[-\left(\frac{\pi}{2}\right)^2 \tau\right] \sin\left(\frac{\pi}{2}\ \xi\right), \tag{15}$$

as $\tau \to \infty$ and

$$\mathcal{M}_c(t) \sim \exp\left[-\left(\frac{\pi}{2}\right)^2 \frac{D}{\delta^2}\ t\right] \frac{8}{\pi^2}, \tag{16}$$

as $t \to \infty$.

Remark 1. At small times, the Fourier series solution becomes inefficient, requiring an increasing number of terms as $t \to 0$. In these cases, a more suitable approach to evaluate the concentration function and its moments is based on an error function expansion. Specifically, the solution of (6) can be expressed as

$$c(\xi, \tau) = \sum_{n=-\infty}^{+\infty} \left[\frac{1}{2} \operatorname{Erf}\left(\frac{\xi - 4n}{\sqrt{4\tau}}\right) - \operatorname{Erf}\left(\frac{\xi - 2 - 4n}{\sqrt{4\tau}}\right) + \frac{1}{2} \operatorname{Erf}\left(\frac{\xi - 4 - 4n}{\sqrt{4\tau}}\right)\right], \tag{17}$$

and its first moment is given by

$$\mathcal{M}_c(\tau) = \sum_{n=-\infty}^{+\infty} \left(\frac{I_1}{2} - I_2 + \frac{I_3}{2}\right), \tag{18}$$

where

$$I_1 = \sqrt{\frac{4\tau}{\pi}} \left[\exp\left(-\frac{(1-4n)^2}{4\tau}\right) - \exp\left(-\frac{4n^2}{\tau}\right)\right] - 4n \operatorname{Erf}\left(\frac{2n}{\sqrt{\tau}}\right) + (1-4n) \operatorname{Erf}\left(\frac{1-4n}{\sqrt{4\tau}}\right),$$

$$I_2 = \sqrt{\frac{4\tau}{\pi}} \left[\exp\left(-\frac{(1+4n)^2}{4\tau}\right) - \exp\left(-\frac{(1+2n)^2}{\tau}\right)\right] - 2(1+2n) \operatorname{Erf}\left(\frac{1+2n}{\sqrt{\tau}}\right) + (1+4n) \operatorname{Erf}\left(\frac{1+4n}{\sqrt{4\tau}}\right),$$

$$I_3 = \sqrt{\frac{4\tau}{\pi}}\left[\exp\left(-\frac{(3+4n)^2}{4\tau}\right) - \exp\left(-\frac{4(1+n)^2}{\tau}\right)\right] - 4(1+n)\,\mathrm{Erf}\left(\frac{2(1+n)}{\sqrt{\tau}}\right) + (3+4n)\,\mathrm{Erf}\left(\frac{(3+4n)}{\sqrt{4\tau}}\right).$$

Remark 2. The Fourier-sine and error function expansions given above are analytical representations of the same solution, and so yield identical results. We verified that this is indeed the case, based on direct evaluation of the corresponding expressions. This was conducted at small, intermediate, and large times, and the exercise reflected the rapid convergence of the error function expansion at small times, and of the sine series representation at large times.

Remark 3. In the computations below, we rely exclusively on the Fourier-sine representation, particularly on the mixed measure representation given by (11). This is primarily driven by computational savings associated with the rapid convergence of the Fourier expansion for times that are of the order of the mixing time scale or larger. The corresponding efficiency is significant, since multiple evaluations of the analytical representation of the mixed measure are required to perform the Bayesian inference analysis.

3. Atomistic system. In this section, we provide a brief outline of the MD computations and, then, discuss the development of a mixing measure approach to quantify the amount of atomic mixing in the MD system.

3.1. MD computations. In order to extract the diffusivity in Ni/Al bilayers within a Bayesian inference framework, our first goal is to build a database of MD computations. The computations, performed with LAMMPS [32] (lammps.sandia.gov), are based on the embedded atom method (EAM) [33, 34, 35]. In this model, the total energy E_{total} is expressed as

$$E_{total} = \frac{1}{2}\sum_{i,j\neq i}\Phi_{ij}(r_{ij}) + \sum_i F_i(\bar{\rho}_i), \tag{19}$$

where $\Phi_{ij}(r_{ij})$ is the two-body pairwise potential between atoms i and j separated by a distance r_{ij}, and F_i is the embedding energy, i.e., the energy to insert an atom i into an electron gas of density $\bar{\rho}_i$. The host electron density is calculated assuming that the electron density in the vicinity of atom i can be approximated as a linear superposition of isotropic contributions from individual atoms according to

$$\bar{\rho}_i = \sum_{j\neq i}\rho_j(r_{ij}), \tag{20}$$

where $\rho_i(r_{ij})$ is the electron density of atom i due to atom j. The pairwise interaction, electron density, and embedding energy depend on the chemical species. The embedding function is assumed to be a purely local function of the electron distribution. In the present work, we consider the EAM potential developed by Mishin [36], which was parametrized by using experimental properties combined with a large set of first-principles data. This potential provides an accurate description of a large variety of properties, such as phonon frequencies, thermal expansion, diffusion, and equations of state for Al, Ni, and their alloys. Note that although this EAM potential reproduces many aspects of the Ni/Al phase diagram, the computed formation energies of Al-rich compounds lie above the NiAl-Al line and are not present as stable intermetallic phases, in contradiction to the experimental phase diagram.

Periodic boundary conditions are implemented in all three directions to approximate the effect of an infinite array of layers in the z direction that extend without bound in the x and y directions. The simulation domain is of size $80 \times 31 \times 31$ Å in the x, y, and z directions, respectively. Initially, the system is at room temperature ($T = 300$ K) and comprised of compositionally pure layers, with Ni atoms occupying the region $0 \leq x \leq 31$ Å, and Al atoms occupying the region $31 \leq x \leq 80$ Å; see Figure 2. The initial Ni/Al bilayer used in all of the simulations presented here is composed of 2916 Ni and 3072 Al in an FCC lattice, arranged with [001] along z direction. The ratio of Ni and Al atoms is approximately 0.95; the composition is thus close to 1 : 1.

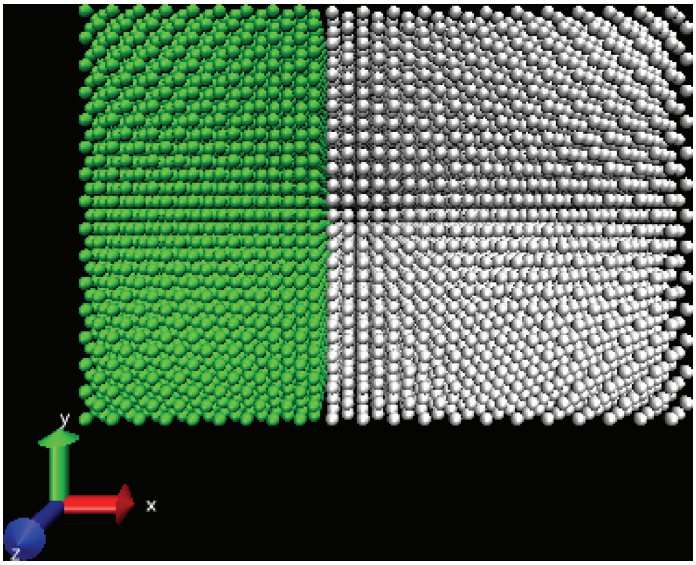


FIG. 2. *Snapshot of the Ni/Al bilayer at time $t = 0$, showing the atoms distribution: Ni (green) and Al (white).*

The MD computation is divided in two main parts. First, the system is quickly heated from $T = 300$ K to the desired temperature, $T = 1500$ K, at constant zero pressure using a time step $\Delta t = 0.005$ ps. For this rapid heating stage, 20000 integration steps are performed. Using the same time step, the whole system is then simulated for 300000 integration steps under isobaric-isothermal ensemble control (NPT), setting $P = 0$ Pa and $T = 1500$ K. The pressure and temperature of the system are maintained in equilibrium with a heat and mechanical reservoir using a Nose–Hoover/Parrinello–Rahman formalism [37, 38, 39, 40, 41]. Dynamic feedback between these reservoirs and the system allows the volume and the kinetic energy to fluctuate about the desired mean values as expected in an equilibrium system. The barostatting and thermostatting time scales were 2 ps. Attention is focused on the second phase of the simulation, during which the mixing occurs under isothermal conditions.

In order to quantify potential effects of variability in the MD database, vacancies are initially introduced in the system by randomly deleting 0.5% of both Ni and Al atoms. By drawing different realizations from a pseudorandom number generator, 80 different initial configurations are generated starting from the same "nominal" system with no vacancies. For each realization thus obtained, the behavior of the system is

determined following the procedure outlined above. Though the level of vacancies is admittedly large, the computed behavior of the systems with vacancies, as determined using the integral analysis below, was found to be very close to that of the nominal system. Thus, for the present conditions, the introduction of vacancies enabled us to obtain a database sufficiently large for the purpose of the Bayesian inference of the atomic diffusivity, without substantial impact on the integral features of the mixing process.

3.2. Mixing measure for the atomistic system. The main purpose of the present analysis is to define a mixing measure that can be used to quantify the amount of atomic mixing in the MD computations, and that can be suitably contrasted to the mixing measure used in the Fickian continuum model, namely the normalized first moment, $\mathcal{M}_c(t)$, of the scalar concentration $c(x,t)$; see (11). To this end, we start by defining the instantaneous CDFs of each constituent in the MD system according to

$$\mathcal{C}_i(x,t) \equiv \frac{n_i(x,t)}{N_i} \qquad \text{for} \quad i = 1,2, \tag{21}$$

where $n_i(x,t)$ refers to the number of atoms of i-th type at time t whose abscissae are smaller than x, and N_i is the total number of atoms of the same type in the entire domain. For the present NPT simulations, the N_i's are constant. The CDFs $\mathcal{C}_1$ and $\mathcal{C}_2$ can be readily estimated from MD computations, based on the instantaneous coordinates of the atoms. Based on these CDFs, we define the following integral measure:

$$\mathcal{M}_d(t) \equiv \frac{1}{2}\int_0^\lambda \left|\frac{\partial \mathcal{C}_1(x,t)}{\partial x} - \frac{1}{\lambda}\right| + \left|\frac{\partial \mathcal{C}_2(x,t)}{\partial x} - \frac{1}{\lambda}\right| dx, \tag{22}$$

where λ is the bilayer thickness; see Figure 1.

For an Ni/Al bilayer comprised of compositionally pure layers at $t = 0$, as in the present computations, the Ni atoms (type 1) are initially located in the region $0 \le x < \beta$, whereas the Al atoms (type 2) are confined in the region $\beta \le x \le \lambda$, as shown in Figure 2. With this convention, the half-thickness δ of the aluminum layer is given by

$$\delta = \frac{\lambda - \beta}{2}. \tag{23}$$

At $t = 0$, we then have

$$\frac{\partial \mathcal{C}_1}{\partial x}(x, t=0) = \begin{cases} \dfrac{1}{\beta} & 0 \le x < \beta, \\ 0 & \beta \le x \le \lambda, \end{cases} \tag{24}$$

whereas

$$\frac{\partial \mathcal{C}_2}{\partial x}(x, t=0) = \begin{cases} 0 & 0 \le x < \beta, \\ \dfrac{1}{\lambda - \beta} & \beta \le x \le \lambda. \end{cases} \tag{25}$$

As an example, Figure 3(a) shows the instantaneous profiles at $t = 0$ of $\mathcal{C}_1$ and $\mathcal{C}_2$ for the Ni/Al bilayer. The corresponding value of the mixing measure $\mathcal{M}_d(t = 0)$ can be obtained from (22)

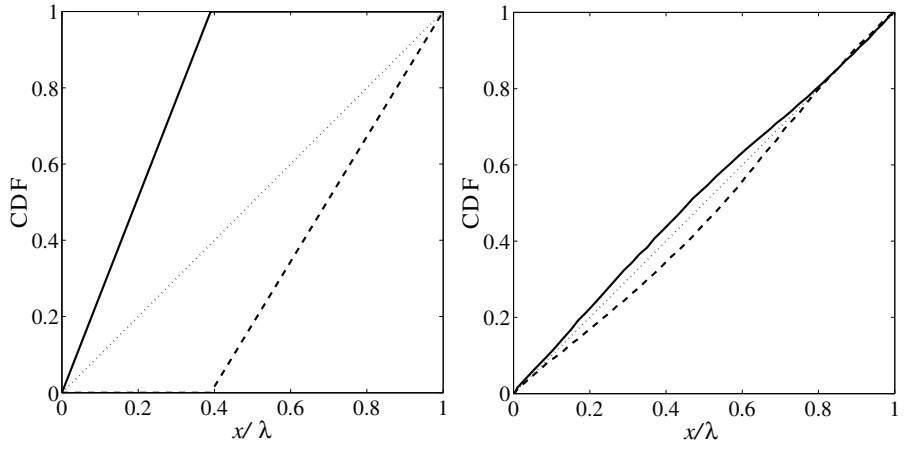


FIG. 3. *CDFs of nickel (solid-line) and aluminum (dashed-line) at* 1500 *K, computed at time* $t = 0$ (a) *and at the instant when* $\mathcal{M}_d = 0.1$ (b). *The dotted line* $y = x$ *corresponds to the asymptotic limit of the CDFs, i.e., when the system is completely mixed.*

$$\begin{aligned}\mathcal{M}_d(t=0) &= \frac{1}{2}\int_0^{\beta}\left|\frac{\partial \mathcal{C}_1(x,t=0)}{\partial x} - \frac{1}{\lambda}\right| dx + \frac{1}{2}\int_0^{\beta}\left|\frac{\partial \mathcal{C}_2(x,t=0)}{\partial x} - \frac{1}{\lambda}\right| dx \\ &+ \frac{1}{2}\int_{\beta}^{\lambda}\left|\frac{\partial \mathcal{C}_1(x,t=0)}{\partial x} - \frac{1}{\lambda}\right| dx + \frac{1}{2}\int_{\beta}^{\lambda}\left|\frac{\partial \mathcal{C}_2(x,t=0)}{\partial x} - \frac{1}{\lambda}\right| dx, \end{aligned} \tag{26}$$

which, upon substitution of (24) and (25), leads to

$$\begin{aligned}\mathcal{M}_d(t=0) &= \frac{1}{2}\left[\beta\left(\frac{1}{\beta} - \frac{1}{\lambda}\right) + \frac{\beta}{\lambda} + \frac{\lambda-\beta}{\lambda} + (\lambda-\beta)\left(\frac{1}{\lambda-\beta} - \frac{1}{\lambda}\right)\right] \\ &= \frac{1}{2}\left[1 - \frac{\beta}{\lambda} + \frac{\beta}{\lambda} + 1 - \frac{\beta}{\lambda} + 1 - 1 + \frac{\beta}{\lambda}\right] \\ &= \frac{1}{2}\left[1+1+1-1\right] = 1. \end{aligned} \tag{27}$$

Hence, the initial values $\mathcal{M}_d(t=0)$ and $\mathcal{M}_c(t=0)$ (see (13)) are identical.

As time progresses, the system tends towards a state where atoms of different type are uniformly distributed over the domain or, in other words, an ideally mixed system; see Figure 3(b). (Due to intrinsic noise in the MD system, however, fluctuations are always observed, even at large times.) For such an idealized state, the CDFs $\mathcal{C}_1$ and $\mathcal{C}_2$ coincide and are given by

$$\mathcal{C}_1(x, t\to\infty) = \mathcal{C}_2(x, t\to\infty) = \frac{1}{\lambda}\, x \qquad \text{for} \quad 0 \le x \le \lambda\ , \tag{28}$$

which immediately yields

$$\mathcal{M}_d(t\to\infty) = 0. \tag{29}$$

Thus, the large-time asymptotic values of $\mathcal{M}_d$ and $\mathcal{M}_c$ also coincide.

From the discussion above, we note that $\mathcal{M}_d$ is a symmetrized first moment of the CDFs that starts at $\mathcal{M}_d = 1$ for an unmixed state and asymptotically tends

to 0 for a well-mixed system, whereas $\mathcal{M}_c$ is a first moment of the concentration profile, starting at $\mathcal{M}_c = 1$ and tending towards 0 as $t \to \infty$. Consequently, a suitable inference strategy can be based on "comparing" analytical estimates of $\mathcal{M}_c(t)$ for a given D, with computed MD-based realizations of $\mathcal{M}_d(t)$. A simplified approach consistent with this strategy is developed below.

4. Bayesian inference. A simplified approach to the inference of the diffusion coefficient is adopted that relies on a single scalar observable rather than the entire history of the mixing measure, $\mathcal{M}_d(t)$. To this end, we define as "conversion time" the value of t/δ^2 at which the mixing measure $\mathcal{M}_d$ first reaches a specified threshold. For instance, the 90% conversion time corresponds to the value of t/δ^2 at which $\mathcal{M}_d = 0.1$. Let $\tilde{t}_d$ be the conversion time computed for a single realization of the atomistic system. A likelihood function can then be constructed from the following relationship:

$$\tilde{t}_d = \tilde{t}_c(D) + \epsilon, \tag{30}$$

where $\tilde{t}_c(D)$ is the corresponding conversion time computed with the continuum model, which is a function of the diffusion coefficient D. The discrepancy between the continuum prediction and the realizations of the MD system is captured by the random variable ϵ. With N independent realizations of the MD computations, we introduce the likelihood function,

$$L(D) \equiv p\left(\{\tilde{t}_d^i\}_{i=1}^N \,|\, D\right) = \prod_{i=1}^{N} p_\epsilon\left(\tilde{t}_d^i - \tilde{t}_c(D)\right), \tag{31}$$

where p_ϵ represents the probability density associated to the random variable ϵ.

A key step in every inference analysis [42] concerns the representation of the distribution of ϵ. Based on central limit theorem arguments, we assume ϵ to be Gaussian with mean zero and variance σ^2; in short $\epsilon \sim N(0, \sigma^2)$. We will verify this distributional hypothesis both a priori and using posterior predictive checks. Also, we will not prescribe a fixed value for the variance σ^2, but we will treat it as a hyperparameter and jointly infer σ^2 and D. Using Bayes' theorem, the joint posterior distribution of these parameters can be written as

$$p\left(D, \sigma^2 | \{\tilde{t}_d^i\}_{i=1}^N\right) = \frac{p\left(\{\tilde{t}_d^i\}_{i=1}^N | D, \sigma^2\right) \hat{p}(D)\hat{p}(\sigma^2)}{A}, \tag{32}$$

where A is a normalization coefficient. Evaluating the posterior probability density requires the definition of the prior densities $\hat{p}(D)$ and $\hat{p}(\sigma^2)$. The priors represent knowledge about the parameters before analyzing the MD simulations; see [42] for more details. In this case, we assume a priori independence. Since sharp priors are generally not available, and since a large MD database can be obtained in the present isothermal setting, we rely on relatively noninformative priors. A suitable and convenient choice is a Jeffreys prior for both D and σ^2, expressed in the following form [42]:

$$\hat{p}(D) = \frac{1}{D} \quad \text{and} \quad \hat{p}(\sigma^2) = \frac{1}{\sigma^2}. \tag{33}$$

The problem now reduces to simulating from the posterior density (32). In general, when the space of the unknown parameters is multidimensional, a suitable computational strategy is based on Markov chain Monte Carlo (MCMC). However, since our problem is two-dimensional, the joint posterior (32) can be readily evaluated on a grid of D and σ^2 values. One can then marginalize over σ^2 to obtain the quantity of interest, i.e., the posterior distribution of D alone. In the analysis below, we perform and compare both approaches for the purpose of verifying predictions.

5. Results. In this section, we discuss the results obtained using the inference approach outlined above. In subsection 5.1, we examine the MD database, in particular verifying that the noise in the MD computations is normally distributed. Inference results are then presented in subsection 5.2, and *a posteriori* verification of the predictions is conducted in 5.3. In subsection 5.4, we briefly explore the possibility of extracting the mean diffusivity based on the analysis of the mean-square displacements at large time. A discussion is finally provided in subsection 5.5 in light of available experimental correlations.

5.1. A priori analysis of MD predictions. In the simplified analysis of section 4, the discrepancy between the continuum prediction and the MD realizations was represented in terms of a random variable ϵ; see (30). The variable ϵ was assumed to be normally distributed. In order to verify this hypothesis, we perform a statistical analysis based on computing the quantiles of the MD realizations of a given conversion time, $\tilde{t}_d$.

Figure 4 shows the evolution of the integral measure for 50 realizations of the atomistic system. Specifically, $\mathcal{M}_d$ is plotted against the reduced time scale t/δ^2. Note that in the early stages of the computations, a sharp transient is observed, associated with an abrupt drop in $\mathcal{M}_d$. This appears analogous to observations made in simulations of intermixing in nanoparticles, particularly for systems involving size mismatch [43]. The sharp transient is of short duration, and accounts for a small fraction of the initial heating stage, $t/\delta^2 \leq 0.16$. Subsequently, the mixing measure evolves smoothly, decaying monotonically in a manner that is reminiscent of a Fickian process.

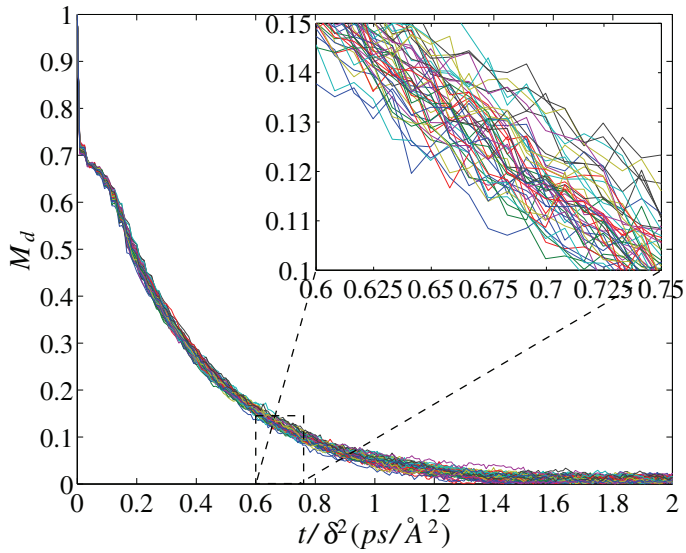


FIG. 4. *Evolution of the integral measure $\mathcal{M}_d$ as a function of t/δ^2 computed for 50 realizations of the Ni/Al bilayer of size $8 \times 3.1 \times 3.1$ nm. Each realization is obtained by randomly imposing a 0.5% percentage of vacancies to the system.*

Figure 5(a)–(d) shows the quantile-quantile (Q-Q) plots versus normally distributed data, of $\{\tilde{t}_d^i\}_{i=1}^{80}$, computed at $\mathcal{M}_d = 0.3$, 0.2, 0.1, and 0.05, i.e., the 70, 80, 90, and 95% conversion times. This analysis allows us to assess whether the

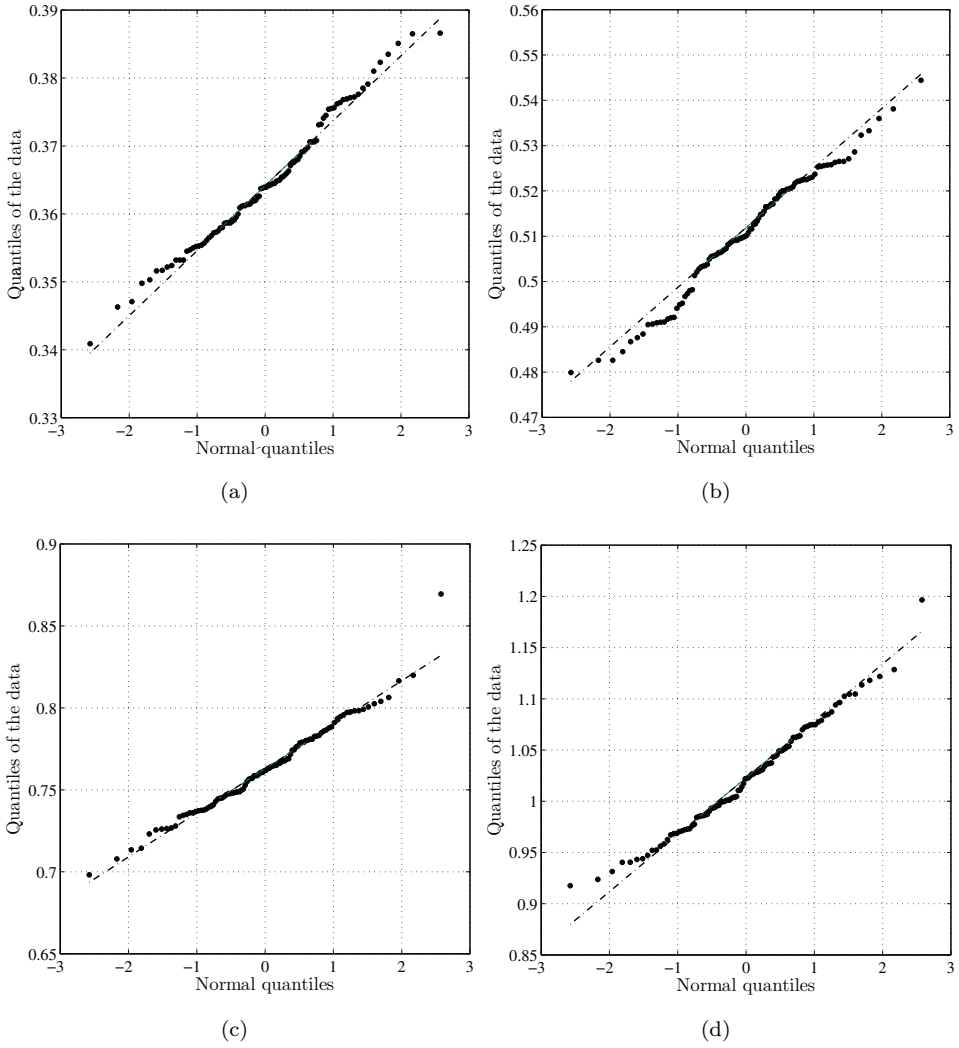
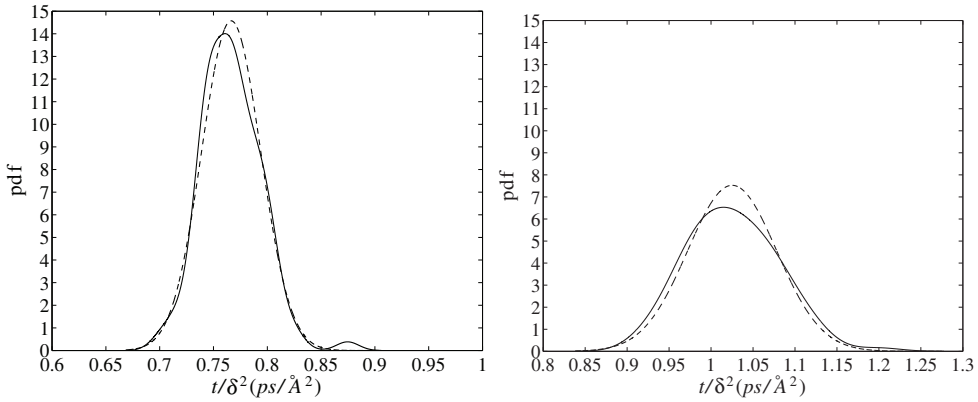


FIG. 5. *Q-Q plots versus normal distributed data of the MD realizations* $\{\tilde{t}_d^i\}_{i=1}^{80}$*, computed at* $\mathcal{M}_d = 0.3$ (a)*,* 0.2 (b)*,* 0.1 (c)*, and* 0.05 (d)*. The superimposed dash-dotted line helps evaluate the linearity of the data.*

MD data follow a normal distribution. Specifically, the computed curves would be linear when the data are normally distributed, and would exhibit a different trend when this is not the case. The results indicate that as $\mathcal{M}_d$ decreases from 0.3 to 0.05, i.e., as mixing increases, the hypothesis that the MD data are normally distributed becomes increasingly more accurate. For panels (a) and (b) ($\mathcal{M}_d = 0.3$ and 0.2, respectively) the results show that the quantiles substantially deviate from the theoretical prediction, represented by the dash-dotted line. On the other hand, the quantiles shown on panels (c) and (d) ($\mathcal{M}_d = 0.1$ and 0.05, respectively) follow closely the idealized Gaussian curve, with only few outliers at the high and low ends of the range. Consequently, a normal distribution appears to be a valid assumption when $\mathcal{M}_d = 0.1$ and 0.05, whereas it is not appropriate at earlier times, i.e., when $\mathcal{M}_d > 0.1$.

We further examine the hypothesis that the noise is normally distributed by constructing the probability density of the MD realizations $\{\tilde{t}_d^i\}_{i=1}^{80}$, and comparing the results to a Gaussian fit. Figure 6(a)–(b) shows the plots of the probability distributions computed via kernel density estimation; see, e.g., [44, 45, 46, 47], of the MD realizations $\{\tilde{t}_d^i\}_{i=1}^{80}$ at $\mathcal{M}_d = 0.1$ and 0.05, superimposed to the fitted Gaussian model. As expected, a normal distribution provides overall a good description of the density of the MD data for these two values of $\mathcal{M}_d$. However, some differences still arise. In particular, for $\mathcal{M}_d = 0.1$, the probability distribution is slightly asymmetric, though it has nearly the same variance as Gaussian fit; see Figure 6(a). For $\mathcal{M}_d = 0.05$ on the other hand, the computed distribution closely follows the Gaussian fit, but has a slightly larger variance.



(a) (b)

FIG. 6. *Panels* (a) *and* (b) *show the comparison between the probability distribution (solid-line) obtained via kernel density estimation of the data* $\{\tilde{t}_d^i\}_{i=1}^{80}$ *and a Gaussian fit (dashed-line), computed at* $\mathcal{M}_d = 0.1$ *and* 0.05*, respectively.*

5.2. Inferred diffusivity. Figure 7(a)–(d) shows the contour plots of the joint posterior distribution (32) of the diffusion coefficient, D, and noise variance, σ^2, computed using $N = 20, 40, 60$, and 80 realizations of the 95% conversion time. Recall that each realization of the atomistic system corresponds to a different randomly chosen distribution of the vacancies. As the number of realizations increases from $N = 20$ to $N = 80$, the following two observations can be made. First, the joint posterior becomes narrower along the D axis, meaning that the variance of the diffusion coefficient D decreases, reaching its minimum value when $N = 80$. Second, the mode of the distribution varies little with respect to the D axis, whereas it shifts substantially along the σ^2 axis. Note also that the distribution is nearly symmetric with respect to D, whereas it is slightly right-skewed with respect to σ^2.

The posterior distribution of the diffusivity, $p\left(D|\{\tilde{t}_d^i\}_{i=1}^N\right)$, can be readily obtained by marginalizing the joint distribution over σ^2, namely according to

$$p\left(D|\{\tilde{t}_d^i\}_{i=1}^N\right) = \int_0^\infty p(D, \sigma^2 \,|\{\tilde{t}_d^i\}_{i=1}^N)\, d(\sigma^2).$$

The posterior distribution of σ^2 can be similarly computed by marginalizing the joint distribution over D. Figure 8 (a)–(d) shows the marginalized distribution of D

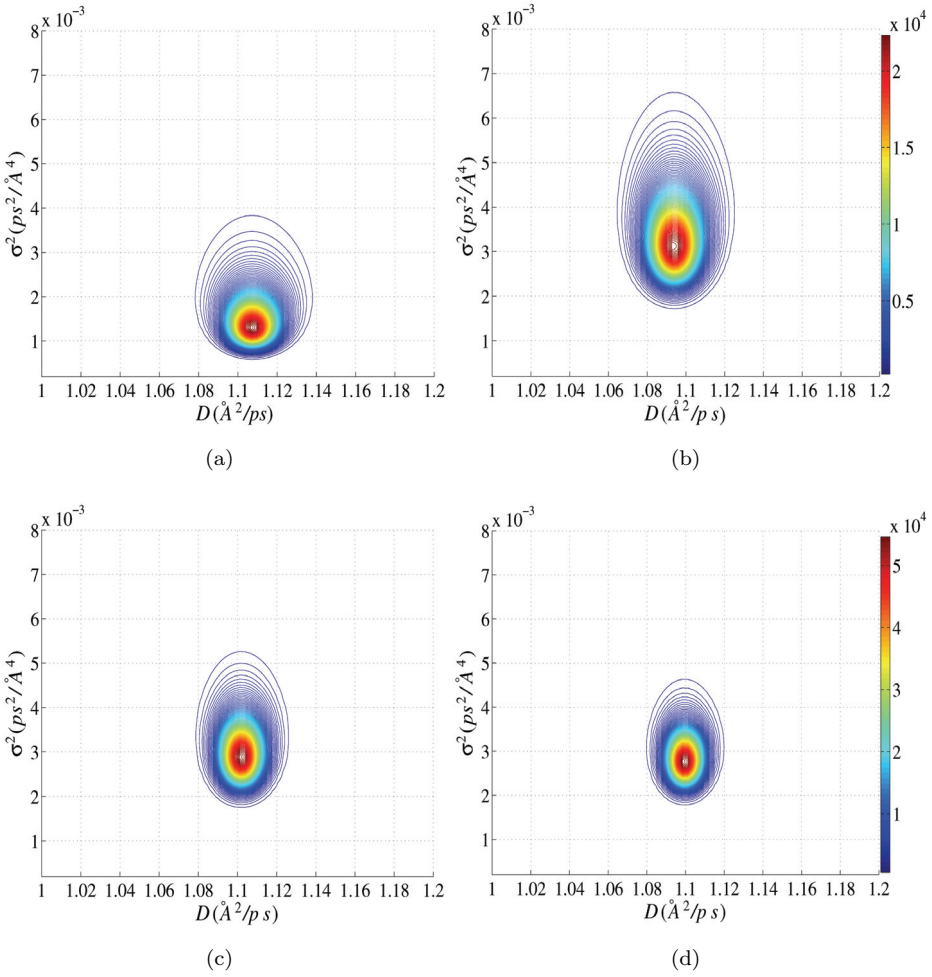


FIG. 7. *Contour plots of the joint posterior of the diffusion coefficient, D, and noise variance, σ^2, based on* 95% *conversion time, using* $N = 20$ (a), 40 (b), 60 (c), *and* 80 (d) *realizations.*

(first row), and σ^2 (second row), based on the joint densities computed using $N = 20$, 40, 60, and 80 realizations of the 90% (first column) and 95% (second column) conversion times.

We first examine the dependence of the distributions of D and σ^2 on the number of realizations N for a given conversion time. Figure 8(a) shows that for 90% conversion time the variance of D monotonically decreases as N increases; in other words the probability distribution becomes narrower. On the other hand, Figure 8(b) shows that for 95% conversion time the variance of D is nearly the same for $N = 20$ and 40 realizations, and suddenly decreases for $N = 60$ and 80. The trend is different if we consider the posterior of σ^2; see Figure 8(c)–(d). Specifically, for fixed conversion time, the variance of the posterior of σ^2 is weakly dependent on N. At 90% conversion time, the variance of the posterior of σ^2 remains nearly the same for $N = 20$, 40, and 60, but the distribution is narrower for $N = 80$; see Figure 8(c). For 95% conversion time, however, the distribution shows highest spread when $N = 40$, and is narrower for the other values; see Figure 8(d).

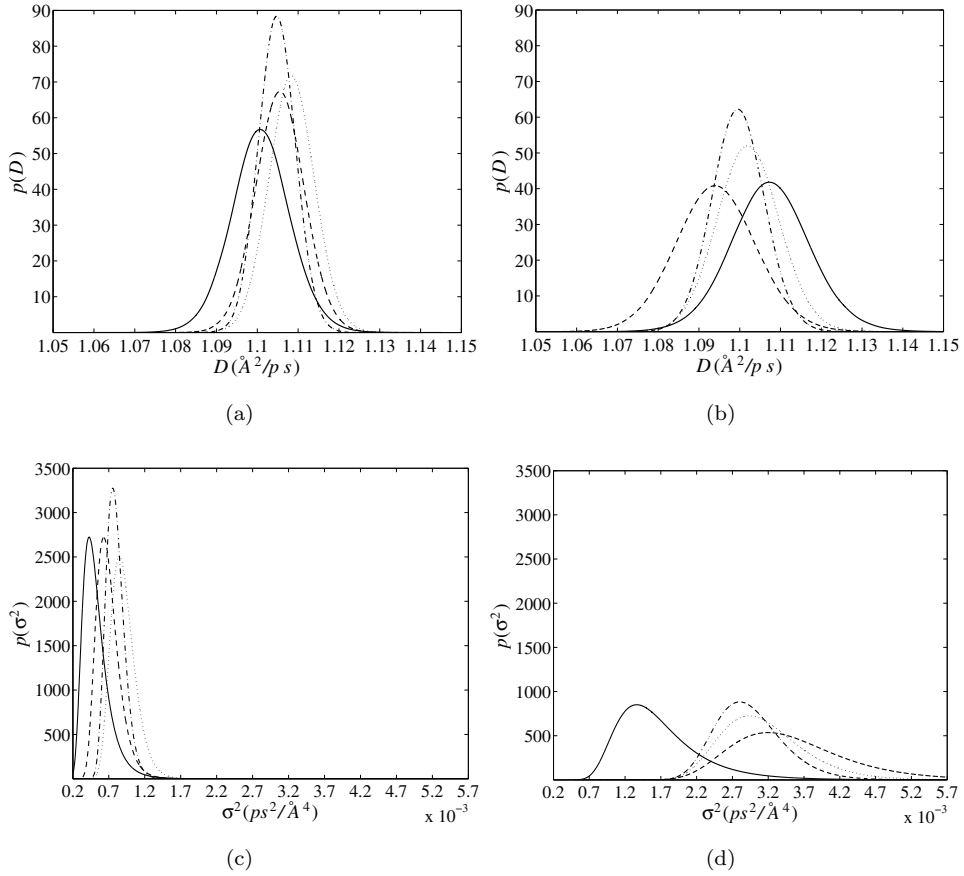


FIG. 8. *Marginalized distributions of the diffusion coefficient D (first row), and noise variance σ^2 (second row), based on* 90% (a), (c) *and* 95% (b), (d) *conversion times, using* $N = 20$ *(solid-line),* 40 *(dashed),* 60 *(dotted), and* 80 *(dash-dotted) realizations.*

We now analyze how the marginalized posteriors of D and σ^2 depend on the conversion time for a given number of realizations. For a given value of N, the distributions of both D and σ^2 become wider as the conversion time increases. This result is expected if we look at the time evolution of the mixing measure $\mathcal{M}_d(t)$; see Figure 4. In particular, one can observe that the curves spread out as the mixing measure tends to zero. However, a major difference is observed between the probability of D and σ^2 when we consider the dependence of the mode on the conversion time. On the one hand, the location of the mode for the posterior of D varies only by few percentage units; see Figure 8(a)–(b). On the other hand, the location of the mode of the posterior of σ^2 substantially varies as the conversion time increases from 90% to 95%; see Figure 8(c)–(d).

As a second step of the inference analysis, we simulate samples from the joint posterior (32) using an MCMC algorithm. This step might appear superfluous for the present two-dimensional problem, in the sense that we have already plotted contours of the posterior density and calculated the marginal densities by direct integration. But the ability to generate samples from the posterior is useful for a variety of subsequent diagnostics and predictions, in particular the posterior predictive checks described in section 5.3.

Leveraging the unimodal nature of the distribution, we employ a "slice sampler" [48, 49, 50], rather than the more common Metropolis–Hastings algorithm. The slice sampler algorithm is based on the concept that in order to sample a random variable, one can sample uniformly from the region under its probability density function. This algorithm is easily applicable to unimodal distributions, as in the present case, where it can exhibit higher efficiency than simple Metropolis updates, due to the ability of adaptively choosing the magnitude of changes made; see [50]. On the contrary, this algorithm becomes more complex to implement, yielding also low efficiency, when applied to multimodal distributions.

Figure 9(a)–(d) shows the scatter plots of the chain, along with the contours of the probability distribution constructed via kernel density estimation; see, e.g., [44, 45, 46, 47]. These results are based on $N = 20$, 40, 60, and 80 realizations of the 95% conversion time. The contour plots in Figure 9 are in very good agreement with the corresponding plots shown in Figure 7(a)–(d), not only from qualitative perspective but also quantitatively. As an example, we note how the location of the mode and the extent of the contours obtained in Figure 7 are nearly exactly reproduced by the slice sampler algorithm.

Figures 10(a)–(b) show, respectively, the plots of the evolution of D and σ^2 obtained from the MCMC for $N = 80$ realizations. Note that, for brevity, the plots for

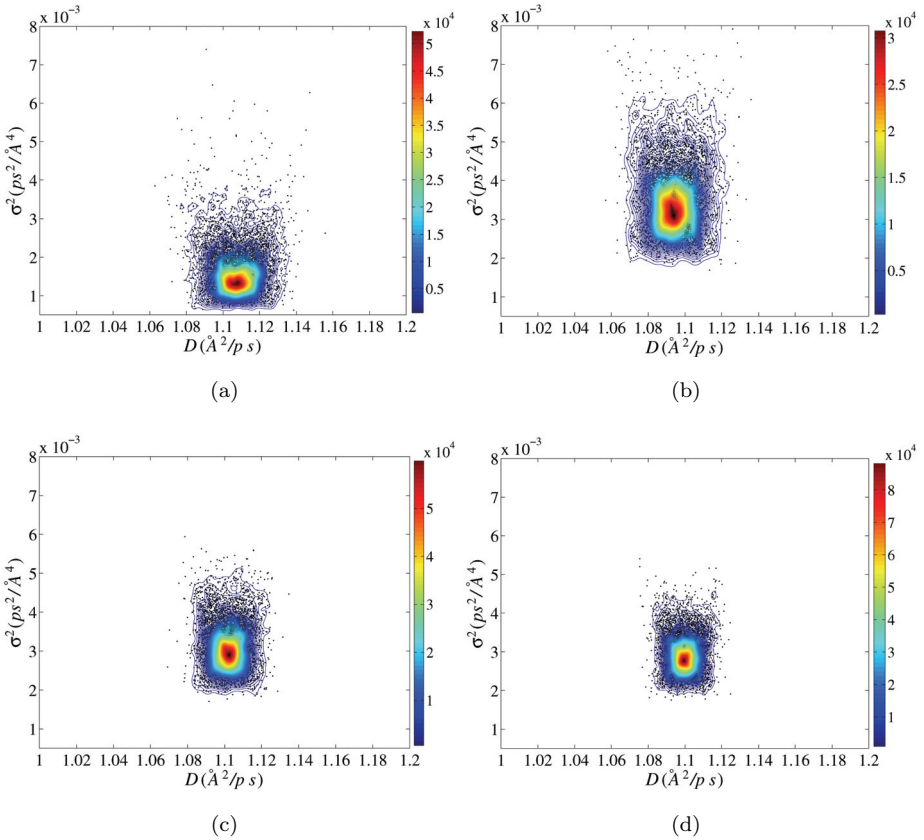


FIG. 9. *Two-dimensional plots showing the Markov chain (dots) along with the contours of the probability distribution obtained via kernel density estimation, obtained for* 95% *conversion time using* $N = 20$ (a), 40 (b), 60 (c), *and* 80 (d) *realizations.*

the cases with $N = 20$, 40, and 60 realizations are omitted because they yield similar results. The evolutions of D and σ^2 are characterized by oscillations whose amplitude is determined by the steps of the chain in the parameter space. As a consequence, the amplitude of these oscillations gives an estimate of the variance of each parameter D and σ^2. To this end, it is important to consider the so-called "burn-in" period. This term is used to identify a set of iterations at the beginning of a Markov chain that is strongly dependent on the starting point, giving an estimate of the time required by the chain to become stationary. Therefore, the choice of the starting point greatly affects the length of the burn-in time. One possibility is to start as close to the center of the distribution as possible, for instance by taking a value in the neighborhood of the mode of the distribution. In general, this is not always possible since the distribution is the objective of the analysis and, therefore, not known a priori. However, one can make use of external sources of information (e.g., experimental data) or optimization methods for maximum a posteriori probability (MAP), to obtain a reasonable estimate of the location of the mode. In our case, however, we take advantage of the known posterior distributions evaluated on the two-dimensional (2D) grid, shown in Figure 7, and, consequently, run the sampling algorithm providing an "optimal" starting point. This choice allows us to reduce to nearly zero the burn-in time. This is

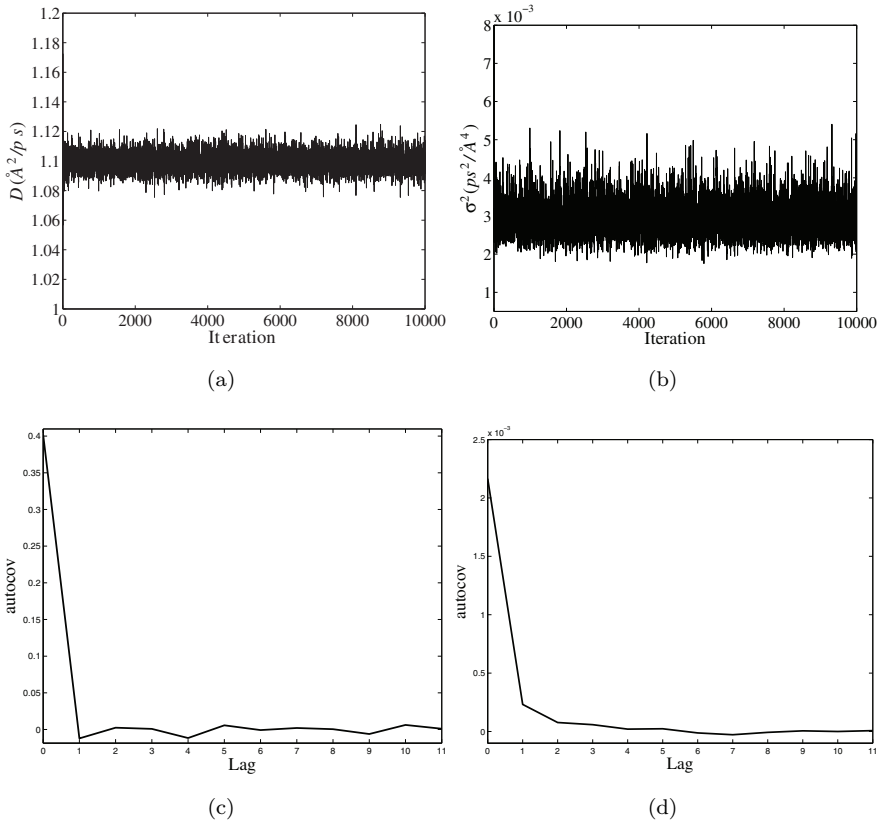


FIG. 10. *Panels* (a) *and* (b) *show the evolution of the x-component, D, and y-component, σ^2, respectively, of the MCMC based on the slice sampler algorithm computed for* 95% *conversion time using $N = 80$ realizations. Panels* (c) *and* (d) *show the evolution of the autocovariance as a function of the lag for each component of the chain, i.e., D and σ^2.*

indeed reflected in the plots of Figure 10(a)–(b), where it is evident that the stationary state is reached already after few iterations.

We finally discuss the mixing property of this Markov chain algorithm. In general, a chain is said to be "poorly mixing" if it is limited to move within a small region of the parameter space for long periods of time. On the contrary, a chain is said "well mixing" when the stochastic space is fully explored. To characterize the mixing property of the MCMC, we compute the autocovariance for each variable, i.e., D and σ^2. Figure 10(c)–(d) shows the plots of the autocovariance as a function of the lag, computed for 95% conversion time using $N = 80$ realizations. In this case, the autocovariance of both D and σ^2 rapidly decays and approaches zero. This trend is consistent with the good mixing visible in Figure 10(a)–(b). Hence, we can conclude that the slice sampler is very efficient at exploring the posterior distribution and the chain is well mixing.

We finally note that the MCMC analysis above was also applied to MD data based on 90% conversion time where similar results were obtained. For brevity, these results are omitted from the present manuscript.

5.3. Posterior predictive checking. In this section, posterior predictive checks [51] are used to assess the consistency of the inference model (both $\mathcal{M}_c$ and the probabilistic assumptions in the likelihood and prior) with the MD data. To this end, we draw samples D_j from the posterior distribution obtained for 95% conversion time with $N = 80$, and compute the corresponding time evolution of the continuum mixing measure, $\mathcal{M}_c(t; D_j)$. This yields an ensemble of paths, which can be contrasted with MD realizations, $\mathcal{M}_d(t)$, that were used to build the probabilistic model. Figure 11 shows the ensemble of paths $\mathcal{M}_c(t; D_j)$, $j = 1, \ldots, 10$, superimposed to 30 MD realizations. Note that since the posterior distribution of D computed using $N = 80$ has a low variance (see Figure 8(b)), all the paths generated from the continuum model are very close to each other.

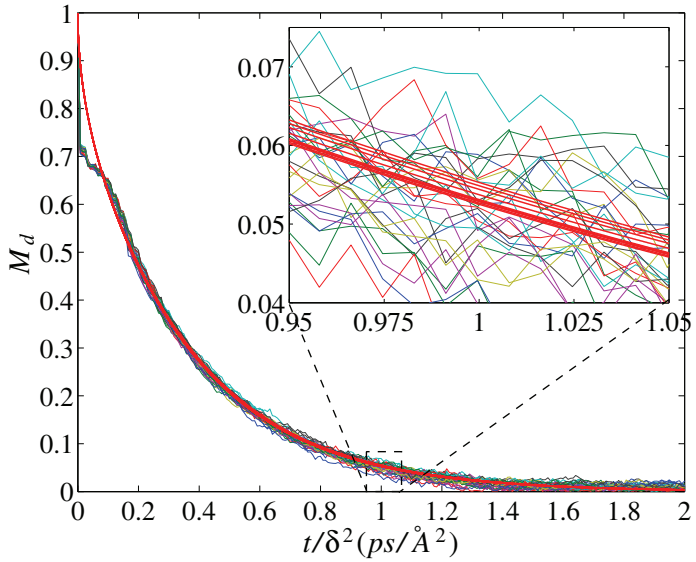
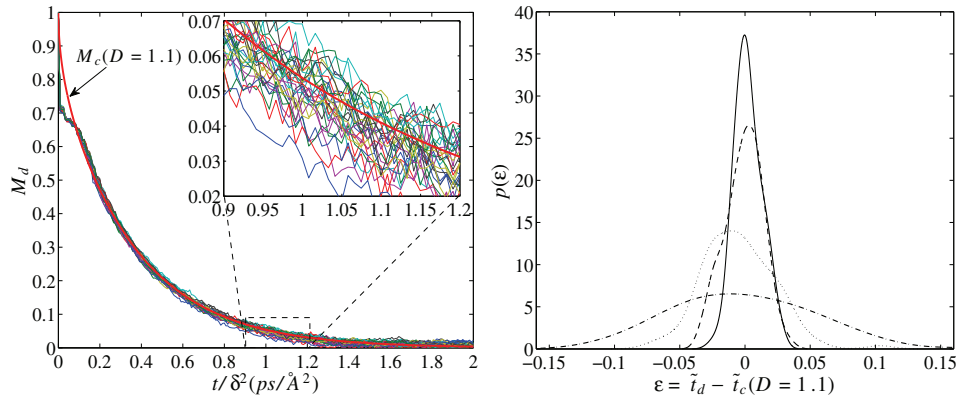


FIG. 11. *Evolution of the mixing measure $\mathcal{M}_d$ as function of t/δ^2 for 30 MD realizations (thin lines) compared to 10 realizations of the continuum model $\mathcal{M}_c$ (red, thick lines) generated by sampling the posterior distribution of D obtained for 95% conversion time and $N = 80$ realizations; see Figure 8(b).*

Examination of the results indicates that in the initial stages of the computations, $t/\delta^2 \leq 0.16$, a noticeable discrepancy occurs between the continuum prediction and the MD realizations. This period corresponds to the initial rapid heating stage, where the Fickian continuum description is clearly not suitable. For $t/\delta^2 > 0.16$, however, the continuum solution closely captures the evolution of the MD-based curves. In particular, good agreement is observed over a substantially large range of variation of the mixing measure, specifically for $\mathcal{M} < 0.7$, and the Fickian representation appears to be suitable to represent the evolution of mixing on time scales of the order of the mixing time scale.

Further analysis of the predictions is performed based on the MAP estimate of D. Figure 12(a) shows the time evolution of the mixing measure $\mathcal{M}_d$ for 30 MD realizations superimposed to the continuum prediction $\mathcal{M}_c(D = 1.1\ \text{Å}^2/\text{ps})$, where the selected value of D corresponds to the MAP estimate of D obtained for 95% conversion time with $N = 80$ (see Figure 8(b)). The profile of the continuum solution is similar to the paths shown in Figure 12 and the evolution of the realizations of the atomistic mixing measure $\mathcal{M}_d$ is well-reproduced. Since the noise, ε, is assumed to have mean zero, we expect the curve for $\mathcal{M}_c(D = 1.1\ \text{Å}^2/\text{ps})$ to fall nearly in the center of the atomistic realizations. This can be verified by constructing and analyzing the empirical distribution of the difference between the MD data $\{\tilde{t}_d^i\}_{i=1}^{80}$, and the continuum solution $\tilde{t}_c(D = 1.1\ \text{Å}^2/\text{ps})$. We perform this analysis for $\mathcal{M}_d = \mathcal{M}_c = 0.3$, 0.2, 0.1, and 0.05, and report the results in Figure 12(b). Note that the variance of the computed distributions increases as $\mathcal{M}$ decreases, and that in all cases the mode is peaked very close to zero.

We finally examine the posterior predictive distribution corresponding to a specific conversion time. A three-step approach is used for this purpose. First, we draw a sample (D_k, σ_k^2) from the joint posterior distribution of D and σ^2 obtained for 95% conversion time with $N = 80$. Second, we compute the continuum solution $\mathcal{M}_c(D_k)$ and find the corresponding 95% conversion time $\tilde{t}_k^c$. Finally, we draw n independently



(a) (b)

FIG. 12. *Panel* (a)*: Evolution of the mixing measure* $\mathcal{M}_d$ *as a function of* t/δ^2 *for* 30 *MD realizations (thin lines) superimposed to the continuum prediction* $\mathcal{M}_c(D = 1.1)$ *(red, thick line), where* $D = 1.1$ *is the MAP estimate of the posterior obtained for* 95% *conversion time and* $N = 80$ *realizations (see Figure* 8(b)*). Panel* (b)*: Probability distribution, obtained via kernel density estimation, of the noise,* ϵ*, between the MD data* $\{\tilde{t}_d^i\}_{i=1}^{80}$*, and the continuum solution* $\tilde{t}_c(D = 1.1)$ *computed at* $\mathcal{M}_d = \mathcal{M}_c = 0.3$ *(solid-line),* 0.2 *(dashed),* 0.1 *(dotted) and* 0.05 *(dash-dotted).*

and identically distributed (i.i.d.) samples $\{\xi^n\}$ from a Gaussian distribution with mean zero and variance σ_k^2, i.e., $\xi^n \sim N(0, \sigma_k^2)$, and generate a set of conversion time samples $\tilde{t}_k^n$ according to

$$\tilde{t}_k^n = \tilde{t}_k^c + \xi^n \qquad \text{for } n = 1, \ldots, 80. \tag{34}$$

We repeat this procedure with four independent draws from the posterior ($k = 1, \ldots, 4$), and compute via kernel density estimation the density of the replicated data $\{\tilde{t}_k^n\}$. Figure 13 compares the probability distributions of t/δ^2 of the MD data at $\mathcal{M}_d = 0.05$ (solid-line) and the distributions obtained as described above. Close agreement between the real and the computed data is observed in the central region, though small discrepancies are noticeable slightly near the tails of the distribution.

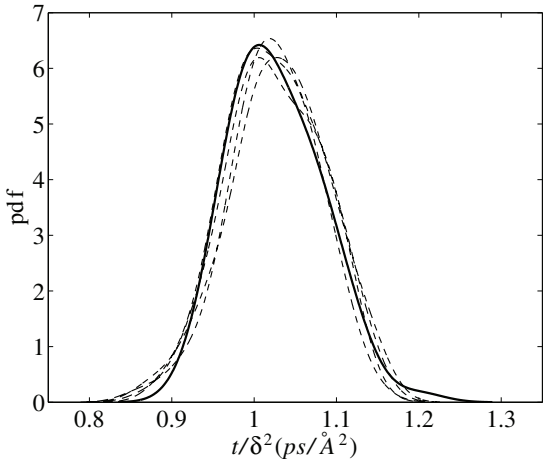


FIG. 13. *Comparison between the probability distribution of the conversion time t/δ^2 for the MD data at $\mathcal{M}_d = 0.05$ (solid-line) and four samples (dashed-lines) obtained from* (34) *for given realizations of D and σ^2.*

Based on the posterior analysis above, it appears that the basic assumptions used in the Bayesian analysis are suitable for the present purpose of inferring the diffusivity.

5.4. Mean-square displacement analysis. In this subsection, we briefly discuss the estimation of diffusivity based on the Einstein–Smoluchowsky diffusion equation (see [52, 53, 54, 55] and references therein):

$$\left\langle |\mathbf{r}(t) - \mathbf{r}(0)|^2 \right\rangle = 2kDt, \tag{35}$$

where t is time, the angle brackets $\langle \cdot \rangle$ represent the ensemble average over all particles, and k is the number of spatial dimensions. When the particle position vector, $r(t)$, is considered, we set $k = 3$ and the displacement corresponds to the Euclidean distance in three-dimensional (3D) space, namely $|\mathbf{r}(t) - \mathbf{r}(0)|^2 = \sum_{i=1}^{3} |r_i(t) - r_i(0)|^2$. On the other hand, we set $k = 1$ when considering displacement along a single space dimension. Equation (35) is typically applied in the limit of large times; i.e., D is extracted based on the asymptotic behavior of the mean-square displacement. We compare the values of D thus determined with those obtained from the Bayesian inference/mixing measure approach above.

Figure 14 shows the evolution of the mean-square displacement (MSD) along with the extracted value of the diffusion constant D, based on the analysis of the x, y, z components of the displacement, as well as the full 3D displacement. In these figures, the evolution of the MSD starts at $t \approx 100$ ps because we only consider the MD data obtained during the isothermal steps of the computation. Note that the values of D extracted from the MSD in Figure 14(a)–(d) are very close to one another. Thus, the MSD analysis indicates that on large time scales diffusion is essentially isotropic.

Comparison of the diffusivity extracted from the MSD computation with the diffusivity inferred from the Bayesian analysis of the mixing measure reveals a large discrepancy. Specifically, the diffusivity derived from the MSD analysis is approximately 0.2 Å^2/ps, whereas the mixing measure analysis yields $D \sim 1.1$ Å^2/ps. Thus, the two estimates differ by nearly an order of magnitude. The occurrence of a discrepancy is not surprising in itself, since the two approaches are based on different definitions of the diffusion coefficient. As discussed in [56], discrepancies of similar nature are observed when comparing estimates for single-particle and collective diffusion coefficients.

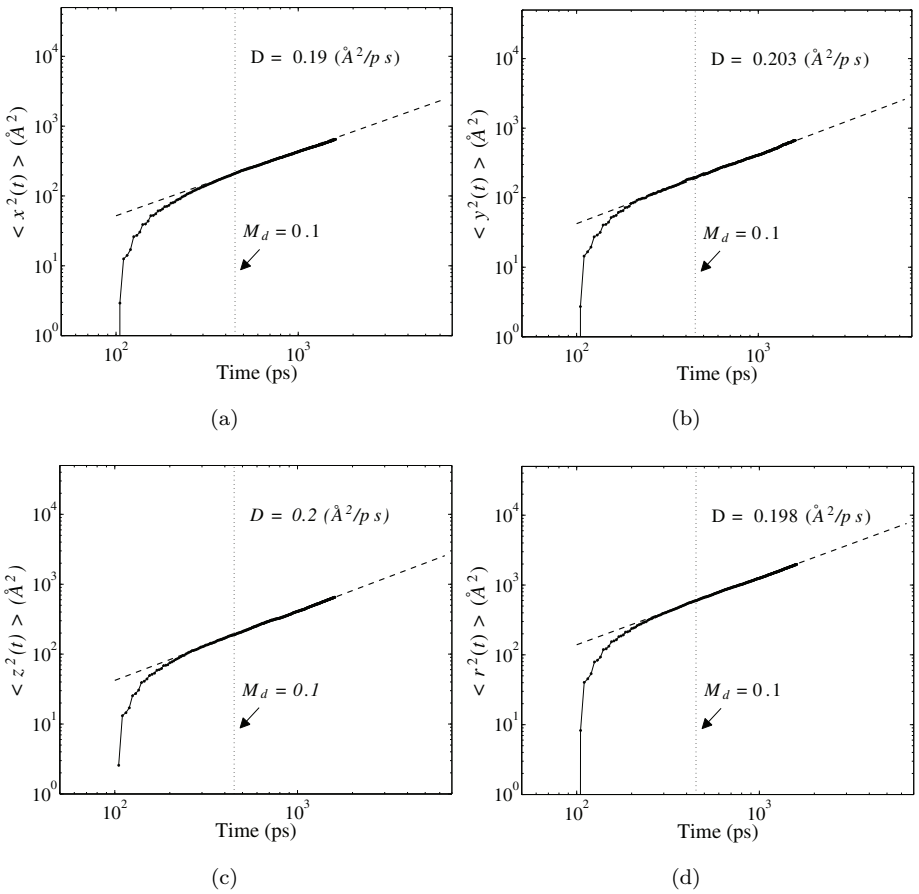


FIG. 14. *Time evolution of the MSD based on the* x (a), y (b), z (c) *component of the displacement and the full displacement* $r = \sqrt{x^2 + y^2 + z^2}$ (d) *for the Ni/Al bilayer at* 1500 *K. The dotted line indicates the time-instant at which* $\mathcal{M}_d = 0.1$*, i.e., a state of nearly complete mixing. The dashed-line is a linear-fit of the large-times region used to compute the diffusion constant,* D.

Insight into the origin of this discrepancy can be gained by further analyzing the results in Figure 14(a)–(d), on which a dotted line is superimposed to indicate the time at which $\mathcal{M}_d = 0.1$, in other words, when mixing is substantially complete. As shown in the figures, $\mathcal{M}_d$ drops to 0.1 by $t \sim 450$ ps. Since the slope of the MSD curves in this interval, $t < 450$ ps, is substantially larger than the asymptotic slope, the latter fails to provide adequate representation of the mixing time scale. Thus, asymptotic analysis of MSD does not appear to be a suitable approach to determine values of the diffusivity that are consistent with the broad features of the present mixing process.

5.5. Discussion. Recall that the Bayesian inference analysis of the mixing measure leads to a MAP estimate of the diffusivity $D \sim 1.1$ Å^2/ps, whereas the MSD analysis leads to $D \approx 0.2$ Å^2/ps. It is instructive to contrast these inferred values of D with experimental results for the diffusion of Ni into Al, and with the Arrhenius fit calibrated from observations of self-propagating fronts. For the sake of clarity, we denote $\hat{D}$ an estimate of the diffusion coefficient based on experimental measurements, whereas D is used to indicate an estimate based on MD computations.

For experimental estimates of the diffusion of Ni into molten Al, we rely on the correlation of Du et al. [57]. The latter is also given in terms of an Arrhenius fit, as in (1), where the activation energy, $E = 26.0$ kJ/mol, and pre-exponential factor, $D_0 = 9.54 \times 10^{-8}$ m^2/s, are estimated based on observed diffusion rates. In the temperature range $952 < T < 1250$ K, the correlation is in very good agreement with the measurements in [31, 58]. Extrapolating this correlation to estimate the diffusivity of Ni into molten Al at $T = 1500$ K, we obtain $\hat{D} \approx 1.2$ Å^2/ps, which we treat as an (approximate) empirical value of the diffusivity of Ni into molten Al at the selected temperature.

Assuming a diffusivity in the form given by (1), Mann et al. [7] determined values of the activation energy and pre-exponent based on experimental measurement of reaction front velocities in Ni/Al multilayers. The best fits resulting from this analysis are $E = 137$ kJ/mol and $D_0 = 2.18^{-6}$ m^2/s. For $T = 1500$ K, the corresponding Arrhenius fit results in $\hat{D} \approx 3.69 \times 10^{-3}$ Å^2/ps.

As value of D representative of the inference analysis, we consider $D = 1.1$ Å^2/ps, which corresponds to the MAP estimate of the posterior distribution computed using $N = 80$ realizations of the 95% conversion time, see Figure 8(b). This MD-based prediction is in very good agreement with the extrapolated empirical value of diffusion of Ni into molten Al, namely $\hat{D} \approx 1.2$ Å^2/ps. The two values differ by nearly 7%. On the contrary, the estimate of D based on the MSD, discussed in subsection 5.4, yields $D \approx 0.2$ Å^2/ps, which differs by more than a factor of 5 from the experimental result for diffusion of Ni into molten Al. As discussed in section 5.4, this is associated with the fact that most of the mixing occurs in the early stages of the computations ($t \leq 450$ ps), during which the asymptotic slope obtained from the Einstein–Smoluchowsky diffusion law (35) fails to provide a reasonable estimate of the corresponding mixing rate. The origin of this discrepancy may be due to the fact that as mixing becomes complete, the isothermal system tends to be dominated by an NiAl phase that is in solid state. Further work is evidently needed to examine the validity of this conjecture, and to investigate the mechanisms leading to the observed drop in the slope of MSD curves at large time.

We further compare our study with the experimental work by Mann et al. [7], which exploited experimental measurements of the velocity of axially propagating flame fronts in Ni/Al bilayers to estimate the activation energy, E, and pre-exponential

factor, D_0, in (1). The resulting values were $E = 137$ kJ/mol and $D_0 = 2.18^{-6}$ m^2/s, yielding a diffusivity at $T = 1500$ K of $\hat{D} \approx 3.69 \times 10^{-3}$ Å^2/ps. This value is three orders of magnitude smaller that the experimental prediction 1.2 Å^2/ps discussed above and two orders of magnitude smaller than the estimate based on MSD described in section 5.4. The discrepancy between the results by Mann et al. [7] and the experimental prediction 1.2 Å^2/ps suggests that atomic mixing at low to moderate temperatures plays a dominant role in determining the front velocity, and that the correlation in [7] may provide a reasonable estimate of mixing rates at these temperatures. In other words, discrepancies that are likely present at high temperatures do not appear to negatively impact the ability of the continuum formalism [7] in predicting front velocities. On the other hand, the present experiences suggest that for some relevant reaction scenarios involving rapid and substantial heating of nanolaminates, for instance due to shock impact or large capacitive discharge, large predictive errors in the reactants' consumption rates may result from the use of the correlation in [7]. Further work is consequently needed to establish limits of validity of the present continuum and reduced formalisms based on correlations such as those in [7], and to devise generalizations having a wider regime of validity.

6. Conclusions. In this work, we characterized the integral features of atomic mixing in Ni/Al nanolaminates using MD computations. Attention was focused on a simplified setting, consisting of a periodic isothermal system at 1500 K. A mixing measure theory to quantify the amount of mixing in an Ni/Al bilayer was developed based on the first moments of the CDFs of the constituents. The MD simulations were applied to generate a database of mixing measures for the same nominal system. The diffusion coefficient was then estimated through a Bayesian inference approach, based on contrasting the MD-measures with the corresponding moment of a dimensionless scalar concentration evolving according to a Fickian process.

For the selected conditions, the Bayesian inference/mixing measure approach yielded a MAP estimate of the diffusivity $D \approx 1.1$ Å^2/ps. This is consistent with the estimate $\hat{D} \approx 1.2$ Å^2/ps, obtained by extrapolating the experimental correlation for diffusion of Ni into molten Al [57]. On the other hand, when the diffusivity was estimated from the large-time evolution of the MSDs, the analysis led to $D \approx 0.2$ Å^2/ps, which differs from the prior two estimates by one order of magnitude. A large discrepancy was also observed between the inferred diffusivity and the semiempirical correlation based on measurements of self-propagating front velocities [7]. Specifically, for a system at 1500 K, the latter resulted in $\hat{D} \approx 3.69 \times 10^{-3}$ Å^2/ps, which is by more than three orders of magnitude smaller than the MD-inferred diffusivity.

Work is currently underway to extend the present formalism to adiabatic systems, which would involve generalizing the present formalism so one can infer a time-dependent diffusivity law. The resulting computations are anticipated to yield additional insight into atomic diffusion rates over suitably wide ranges of temperature and composition, and thus into improved representations suitable for incorporation into reduced reaction models. Results from this ongoing effort will be reported elsewhere.

REFERENCES

[1] M. E. Reiss, C. M. Esber, D. Van Heerden, A. J. Gavens, M. E. Williams, and T. P. Weihs, *Self-propagating formation reactions in Nb/Si multilayers*, Mater. Sci. Eng. A, 261 (1999), pp. 217–222.

[2] E. Ma, C. V. Thompson, L. A. Clevenger, and K. N. Tu, *Self-propagating explosive reactions in Al/Ni multilayer thin-films*, Appl. Phys. Lett., 57 (1990), pp. 1262–1264.
[3] T. S. Dyer, Z. A. Munir, and V. Ruth, *The combustion synthesis of multilayer Ni/Al systems*, Scripta. Metallurgica, Mater., 30 (1994), pp. 1281–1286.
[4] T. P. Weihs, A. J. Gavens, M. E. Reiss, D. Van Heerden, A. E. Draffin, and D. Stansfield, *Chemistry and physics of nanostructured and related non-equilibrium materials*, Minerals, Metals and Materials Society, Warrendale, PA, 1997.
[5] T. P. Weihs, *Self-propagating reactions in multilayer materials*, in Handbook of Thin Film Process Technology, D. A. Glocker and S. I. Shah, eds., Institute of Physics Publishing, London, 1998, Part B, Section F7.
[6] A. S. Rogachev, *Exothermic reaction waves in multilayer nanofilms*, Russ. Chem. Rev., 77 (2008), pp. 21–37.
[7] A. B. Mann, A. J. Gavens, M. E. Reiss, D. Van Heerden, G. Bao, and T. P. Weihs, *Modeling and characterizing the propagation velocity of exothermic reactions in multilayer foils*, J. Appl. Phys., 82 (1997), pp. 1178–1188.
[8] A. J. Gavens, D. Van Heerden, A. B. Mann, M. E. Reiss, and T. P. Weihs, *Effect of intermixing on self-propagating exothermic reactions in Al/Ni nanolaminate foils*, J. Appl. Phys., 87 (2000), pp. 1255–1263.
[9] S. Jayaraman, A. B. Mann, O. M. Knio, G. Bao, and T. P. Weihs, *Numerical modeling of self-propagating exothermic reactions in multilayer systems*, in MRS Symp. Proc. 481, E. Ma, M. Atzmon, P. Bellon, and R. Trivedi, eds., Materials Research Society, Pittsburgh, PA, 1998, p. 563.
[10] S. Jayaraman, A. B. Mann, T. P. Weihs, and O. M. Knio, *A numerical study of unsteady self-propagating reactions in multilayer foils*, in Proceedings of the Twenty-Seventh International Symposium on Combustion, 1998, The Combustion Institute.
[11] S. Jayaraman, O. M. Knio, A. B. Mann, and T. P. Weihs, *Numerical predictions of oscillatory combustion in reactive multilayers*, J. Appl. Phys., 86 (1999), pp. 800–809.
[12] S. Jayaraman, A. B. Mann, M. Reiss, T. P. Weihs, and O. M. Knio, *Numerical study of the effect of heat losses on self-propagating reactions in multilayer foils*, Combust. Flame, 124 (2001), pp. 178–194.
[13] E. Besnoin, S. Cerutti, O. M. Knio, and T. P. Weihs, *Effect of reactant and product melting on self-propagating reactions in multilayer foils*, J. Appl. Phys., 92 (2002), pp. 5474–5481.
[14] R. Knepper, M. R. Snyder, G. Fritz, K. Fisher, O. M. Knio, and T. P. Weihs, *Effect of varying bilayer spacing distribution on reaction heat and velocity in reactive Al/Ni multilayers*, J. Appl. Phys., 105 (2009), 083504.
[15] J. C. Trenkle, J. Wang, T. P. Weihs, and T. C. Hufnagel, *Microstructural study of an oscillatory formation reaction in nanostructured reactive multilayer foils*, Appl. Phys. Lett., 87 (2005), 153108.
[16] J. P. McDonald, V. Carter Hodges, E. D. Jones, Jr., and D. P. Adams, *Direct observation of spinlike reaction fronts in planar energetic multilayer foils*, Appl. Phys. Lett., 94 (2009), 034102.
[17] J. S. Kim, T. LaGrange, B. W. Reed, M. L. Taheri, M. R. Armstrong, W. E. King, N. D. Browning, and G. H. Campbell, *Imaging of transient structures using nanosecond in situ TEM*, Science, 321 (2008), pp. 1472–1475.
[18] I. E. Gunduz, K. Fadenberger, M. Kokonou, C. Rebholz, C. D. Doumanidis, and T. Ando, *Modeling of the self-propagating reactions of nickel and aluminum multilayered foils*, J. Appl. Phys., 105 (2009), 074903.
[19] O. S. Rabinovich, P. S. Grinchuk, M. A. Andreev, and B. B. Khina, *Conditions for combustion synthesis in nanosized Ni/Al films on a substrate*, Phys. B, 392 (2007), pp. 272–280.
[20] M. Salloum and O. M. Knio, *Simulation of reactive nanolaminates using reduced models I. Basic formulation*, Combust. Flame, 157 (2010), pp. 288–295.
[21] M. Salloum and O. M. Knio, *Simulation of reactive nanolaminates using reduced models II. Normal propagation*, Combust. Flame, 157 (2010), pp. 436–445.
[22] M. Salloum and O. M. Knio, *Simulation of reactive nanolaminates using reduced models: III. Ingredients for a general multidimensional formulation*, Combust. Flame, 157 (2010), pp. 1154–1166.
[23] J. F. Zhao, C. Unuvar, U. Anselmi-Tamburini, and Z. A. Munir, *Kinetics of current-enhanced dissolution of nickel in liquid aluminum*, Acta Mater., 55 (2007), pp. 5592–5600.
[24] A. J. Hickl and R. W. Heckel, *Kinetics of current-enhanced dissolution of nickel in liquid aluminum*, Metallurgical Mater. Trans. A, 6 (1975), pp. 431–440.

[25] V. I. Dybkov, *Interaction of iron-nickel alloys with liquid aluminum*, J. Mater. Sci., 28 (1993), pp. 6371–6380.
[26] C.-L. Tsao and S.-W. Chen, *Interfacial reactions in the liquid diffusion couples of mg/ni, al/ni and al/(ni)-al2o3 systems*, J. Mater. Sci., 30 (1995), pp. 5215–5222.
[27] M. M. G. Alemany, C. Rey, and L. J. Gallego, *Computer simulation study of the dynamic properties of liquid ni using the embedded-atom model*, Phys. Rev. B, 58 (1998), pp. 685–693.
[28] M. M. G. Alemany, C. Rey, and L. J. Gallego, *Transport coefficients of liquid transition metals: A computer simulation study using the embedded atom model*, J. Chem. Phys., 109 (1998), pp. 5175–5176.
[29] A. F. Voter and S. P. Chen, *Characterization of Defects in Materials*, Proc. Mater. Res. Soc. 82, Pittsburgh, 1987.
[30] M. Shimoji and T. Itami, *Atomic Transport in Liquid Metals*, Trans Tech Publications, Aedermannsdorf, Switzerland, 1986.
[31] T. Ejima, T. Yamamura, N. Uchida, Y. Matsuzaki, and M. Nikaido, *Impurity diffusion of 4th period solutes (Fe,Co,Ni,Cu and Ga) and homovalent solutes (In and Tl) into molten Aluminum*, J. Japan Inst. Met., 44 (1980), pp. 316–324.
[32] S. J. Plimpton, *Fast parallel algorithms for short-range molecular dynamics*, J. Comput. Phys., 117 (1995), pp. 1–19.
[33] M. S. Daw and M. I. Baskes, *Semiempirical, quantum mechanical calculation of hydrogen embrittlement in metals*, Phys. Rev. Lett., 50 (1983), pp. 1285–1288.
[34] M. S. Daw and M. I. Baskes, *Embedded-atom method: Derivation and application to impurities, surfaces, and other defects in metals*, Phys. Rev. B, 29 (1984), pp. 6443–6453.
[35] M. W. Finnis and J. E. Sinclair, *A simple empirical n-body potential for transition metals*, Philos. Mag. A, 50 (1984), pp. 45–55.
[36] Y. Mishin, *Atomistic modeling of the γ and γ'-phases of the ni-al system*, Acta Mater., 52 (2004), pp. 1451–1467.
[37] M. Parrinello and A. Rahman, *Crystal structure and pair potentials: A molecular-dynamics study*, Phys. Rev. Lett., 45 (1980), pp. 1196–1199.
[38] M. Parrinello and A. Rahman, *Polymorphic transitions in single crystals: A new molecular dynamics method*, J. Appl. Phys., 52 (1981), pp. 7182–7190.
[39] S. Nosé, *A molecular dynamics method for simulations in the canonical ensemble*, Mol. Phys., 52 (1984), pp. 255–268.
[40] S. Nosé, *A unified formulation of the constant temperature molecular dynamics methods*, J. Chem. Phys., 81 (1984), pp. 511–519.
[41] W. G. Hoover, *Canonical dynamics: Equilibrium phase-space distributions*, Phys. Rev. A, 31 (1985), pp. 1695–1697.
[42] D. S. Sivia, *Data Analysis - A Bayesian Tutorial*, Oxford Science Publications, Oxford, 2006.
[43] F. Baletto, C. Mottet, and R. Ferrando, *Growth of three-shell onionlike bimetallic nanoparticles*, Phys. Rev. Lett., 90 (2003), 135504.
[44] E. Parzen, *On estimation of a probability density function and mode*, Ann. Math. Statist, 33 (1962), pp. 1065–1076.
[45] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London, 1986.
[46] B. A. Turlach, *Bandwidth Selection in Kernel Density Estimation: A Review*, Technical report, Institut de Statistique, Belgium, 1993.
[47] L. Wasserman, *All of Statistics: A Concise Course in Statistical Inference*, Springer Texts in Statistics, Berlin, 2005.
[48] G. O. Roberts and J. S. Rosenthal, *Convergence of slice sampler markov chains*, J. R. Stat. Soc. Ser. B Stat. Methodol., 61 (1999), pp. 643–660.
[49] A. Mira and L. Tierney, *Efficiency and convergence properties of slice samplers*, Scand. J. Statist., 29 (2002), pp. 1–12.
[50] Radford M. Neal, *Slice sampling*, Ann. Statist., 31 (2003), pp. 705–767.
[51] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*, Chapman and Hall/CRC, London, 2004.
[52] A. Einstein, *Investigations on the Theory of the Brownian Movement*, Dover, New York, 1956.
[53] M. von Smoluchowski, *Zur kinetischen theorie der brownschen molekularbewegung und der suspensionen*, Ann. Phys., 326 (1906), pp. 756–780. (in German).
[54] R. K. Pathria, *Statistical Mechanics*, Elsevier, New York, 1996.
[55] K. A. Dill and S. Bromberg, *Molecular Driving Forces: Statistical Thermodynamics in Chemistry and Biology*, Garland Science, New York, 2003.

[56] T. Ala-Nissila, R. Ferrando, and S. C. Ying, *Collective and single particle diffusion on surfaces*, Adv. Phys., 51 (2002), pp. 949–1078.

[57] Y. Du, Y. A. Chang, B. Huang, W. Gong, Z. Jin, H. Xu, Z. Yuan, Y. Liu, Y. He, and F. Y. Xie, *Diffusion coefficients of some solutes in fcc and liquid al: Critical evaluation and correlation*, Mater. Sci. Eng. A, 363 (2003), pp. 140–151.

[58] J. P. Praizey, J. P. Garandet, G. Frohberg, A. Griesche, and K. H. Kraatz, *Diffusion experiments in liquid metals - preliminary results*, in Proceedings of the 1st Int. Symp. on Microgravity Research and Applications in Physical Science and Biotechnology, 2000, p. 481.