International Journal for Uncertainty Quantification, 4 (2): 151–170 (2014)

INFERENCE AND UNCERTAINTY PROPAGATION OF ATOMISTICALLY-INFORMED CONTINUUM CONSTITUTIVE LAWS, PART 1: BAYESIAN INFERENCE OF FIXED MODEL FORMS

Maher Salloum[1,*] *& Jeremy Templeton*[2]

[1] *Sandia National Laboratories, 7011 East Avenue, MS 9158, Livermore, California 94550, USA*

[2] *Sandia National Laboratories, 7011 East Avenue, MS 9409, Livermore, California 94550, USA*

Original Manuscript Submitted: 07/04/2013; Final Draft Received: 01/27/2014

Uncertainty quantification techniques have the potential to play an important role in constructing constitutive relationships applicable to nanoscale physics. At these small scales, deviations from laws appropriate at the macroscale arise due to insufficient scale separation between the atomic and continuum length scales, as well as fluctuations due to thermal processes. In this work, we consider the problem of inferring the coefficients of an assumed constitutive model form using atomistic information and propagation of the associated uncertainty. A nanoscale heat transfer problem is taken as the model, and we use a polynomial chaos expansion to represent the thermal conductivity with a linear temperature dependence. A Bayesian inference method is developed to extract the coefficients in this expansion from molecular dynamics (MD) samples at prescribed temperatures. Importantly, the atomistic data are incompatible with the continuum model because of the finite probability of heat flowing in the opposite direction of the temperature gradient; we present a method to account for this in the model. The fidelity and uncertainty in these techniques are then examined. Validation is provided by comparing a continuum Fourier model against a larger all MD simulation representing the true solution.

KEY WORDS: *constitutive law, continuum, atomistic, Bayesian inference, uncertainty, Fourier model, polynomial chaos expansion*

1. INTRODUCTION

The goal of realizing predictive simulation rests upon development of constitutive relationships of sufficient accuracy with quantified uncertainty. An area where existing constitutive models have particular difficulty is nanoscience and engineering because macroscale relationships do not sufficiently resolve the physics at this small scale [1–5]. While continuum mechanics can be derived from molecular mechanics [6, 7] in an average sense, well-known closures from classical mechanics are often inadequate for problems in nanoscience and nanoengineering where atomistic information becomes important. From a mathematical point of view, the challenge is that any nanoscale model must be inherently stochastic and account for the different types of relationships present, e.g., finite length scales which reduce the population of phonons present in a heat transfer problem. Meeting this challenge to develop accurate constitutive models for continuum descriptions of nanosystems has the potential to enable a new generation of devices characterized and optimized by predictive simulation.

To illustrate the origin of the constitutive modeling problem at atomic scales, the following observations are useful. Discretized continuum models always have a smallest resolved length scale associated with the mesh size. While mesh refinement can reduce this error and generate a family of solutions converging to the solution of the originating continuum equations, in reality the applicability of the equations breaks down as the importance of the

*Correspond to Maher Salloum, E-mail: mnsallo@sandia.gov

2152–5080/14/$35.00

discrete atomistic processes becomes increasingly pronounced (e.g., [8]). In contrast, atomistic simulations have a bounding length scale above as only a finite number of atoms can be simulated on a computer. As this length scale is increased, statistical mechanics ever more accurately relates averages over the atomic system to established continuum quantities [9].

One approach to incorporate atomistic information in continuum models is multiscale modeling. There have been many efforts to develop techniques using a multiscale representation, for example in solid mechanics (see the review article by Miller and Tadmor [10]), heat transfer [11, 12], and fluid mechanics [13, 14], although such methodologies are typically used for problems where the region requiring an atomistic description is (loosely speaking) on a lower-dimensional manifold than the overall system of interest, e.g., interfaces, cracks, and defects. In addition, most atomistic-to-continuum methods utilize continuum models based on macroscale behavior like Fourier's law. A notable exception is the concurrent coupling method by Donev et al. [15] to directly simulate Monte Carlo particles and continuum fluid mechanics, which demonstrated that the Landau-Lipfschitz Navier-Stokes equations, in which Brownian fluctuations are retained, are required to obtain an accurate solution in some nanofluidics problems. It is crucial therefore to derive constitutive laws with their associated uncertainties directly informed from atomistic simulations to be appropriate models of the behavior of matter at the nanoscale.

In contrast to multiscale modeling to accurately capture atomistic behavior in small regions, our proposed approach uses uncertainty quantification (UQ) to estimate constitutive models and their uncertainty so the effects of atomic processes can be propagated through much larger regions. (We note there has been recent interest in applying UQ to atomistic and atomistic-to-continuum systems [16–20].) Because of the difficulty in characterizing the mean and fluctuating components of the constitutive relations using existing methods (even central limit theorems may not hold due to small sample sizes), atomistic simulations have to be performed to extract suitable constitutive laws. Figure 1 shows a one-dimensional version of such problem. The continuum simulation is characterized from a large length scale, L_c, down to the smallest resolved length scale representing the mesh size, h. The red dashed line illustrates an atomistic region of appropriate size to characterize the subcontinuum information at that mesh size level. A molecular dynamics (MD) simulation is performed in order to extract suitable constitutive relationships between the problem variables.

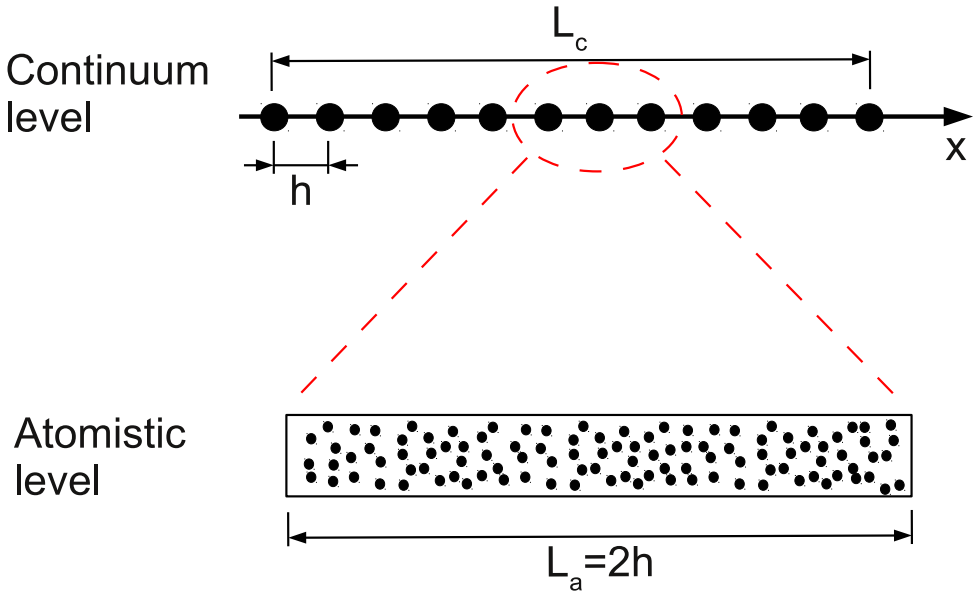

FIG. 1: Schematic showing: (top) a one-dimensional continuum level finite element simulation domain characterized by a length scale L_c, a time scale τ_c, and a mesh size h, and (bottom) an atomistic scale simulation domain characterized by a length scale $L_a = 2h$ and a time scale τ_a, where atomistic information is required to quantify physical phenomena that are beyond the reach of the continuum description.

In this paper, we focus on the study of a one-dimensional heat transfer problem at the nanoscale. This problem offers the advantages of being efficient to simulate on both atomistic and continuum scales while retaining the complexity atomistic-to-continuum coupling because the second law of thermodynamics only holds approximately at the nanoscale [21]. Constitutive law extraction methodologies can also be assessed because the behavior of the atomistically informed continuum model can be compared to the true solution of a large atomic system. We present an approach to extract the relationship between the heat flux and the temperature, as well as temperature gradient, from MD simulations, together with its associated uncertainty.

The controlling assumption in this work is that heat conduction is governed by Fourier's law, specifically, that the heat flux q is the product of the thermal conductivity κ and the temperature gradient ∇T,

$$q = -\kappa(T)\nabla T \tag{1}$$

where κ has a parametric dependency on the temperature T. These parameters are inferred and expressed as polynomial chaos expansions (PCE) due to the facility with which these expansions are capable of propagating uncertainty in continuum simulations [22–26]. Furthermore, PCEs allow a faster convergence in the distribution of an uncertain variable, unlike Monte Carlo methods [25]. The parameters are built using Bayesian inference [27–29], which has been found highly effective in handling different sources of uncertainty including noisy data typically obtained in MD simulations [16, 17]. We extract the heat conduction constitutive law and quantify its associated uncertainty as a function of the spatial- and time-averaging scales and the amount of data used. These are closely related to the time step and spatial discretization sizes in the continuum model. We then propagate the obtained constitutive law into a continuum scale simulation and compare with an equivalent, but fully discrete simulation consisting only of atoms.

This paper is organized as follows. Section 2 overviews the MD and multiscale methods used in this work. The mathematical formulation and specific formulation for the inferred conductivity with specific implementation details are given in Section 3. Results are provided in Section 4 and some concluding thoughts are offered in Section 5.

2. MD SIMULATION

Figure 2 shows a stationary quasi-1D bar simulated by MD as a three-dimensional domain of size $L_a \times w \times w$ in the x, y, and z directions, respectively. MD computations were performed with the LAMMPS [30] atomic simulator. The domain height and width are $w = 3.24$ nm; it is occupied by argon atoms with mass 39.95 g/mol with a FCC lattice constant of $\alpha = 5.405$ Å. The soft particle-particle interaction is modeled by the Lennard-Jones (LJ) pairwise potential ϕ_{ij} [31–33]. For particles i and j separated by a distance r_{ij}, ϕ_{ij} is given by

$$\phi_{ij} = 4\phi_0 \left[\left(\frac{\rho}{r_{ij}} \right)^{12} - \left(\frac{\rho}{r_{ij}} \right)^{6} \right] \tag{2}$$

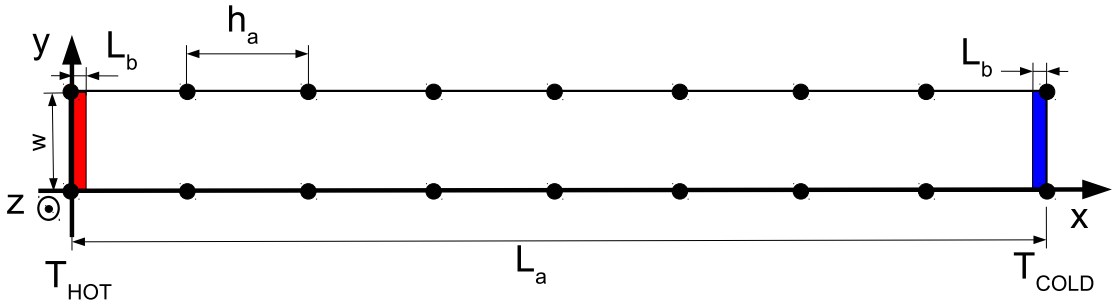

FIG. 2: Schematic showing the MD simulation domain. The temperature is constrained in the red and blue regions to be $T = T_{HOT}$ and $T = T_{COLD}$, respectively. The black dots represent mesh points where the local heat flux, temperature gradient, and temperature are extracted using the formalism of Zimmerman et al. [9].

We set the LJ parameters for argon to $\phi_0 = 0.238$ Kcal/mole and $\rho = 3.405$ Å [31]. Atomic motions are resolved with a time step size of $\Delta t_a = 2$ fs, which was seen to be sufficiently small to yield results independent of the time step in numerical tests. The boundary conditions are periodic in the y and z directions and fixed in the x direction. As the fixed boundary atoms specify a lattice spacing, no energy minimization or prescribed pressure were used in initializing or running the simulation.

This MD simulation is intended to model a quasi-1D heat transfer problem, thus we must impose Dirichlet boundary conditions at $x = 0$ and $x = L_a$. To do so, we use the atomistic-to-continuum formalism originally proposed by Wagner *et al.* [34] to partition the domain using a finite element (FE) mesh. Temperatures are held fixed at the left (T_{HOT}) and right (T_{COLD}) by augmenting the atomic forces from the interatomic potential using a Lagrange multiplier which takes the form of a Gaussian isokinetic thermostat varying on the length scale of the elements (see [35] for further details on performing this type of simulation). It takes about 2.5 ns for the 1D bar to reach a statistical steady state. At this time, short-time averaged samples of the heat flux q (W.m^{-2}), ∇T (K.m^{-1}), and T (K) are collected at the mesh points using the coarse-graining postprocessor of Zimmerman et al. [9] with a localization function ψ characterized by a length scale λ_a. The atomic heat flux is also computed according to the formula in Zimmerman et al. [9] [Eq.(56) therein].

Figure 3 depicts the thermal steady state of the quasi 1D bar. Higher temperatures cause higher atom velocities and more pronounced fluctuations resulting in higher variances in their statistical distributions. Thus, the noise level slightly decreases when the temperature decreases from T_{HOT} to T_{COLD} which, for illustration, were chosen as 60 K and 40 K, respectively. The thermal conductivity is computed during the MD simulation as a finite difference approximation taken on a coarse-grained continuum field, thus avoiding the need to take the derivative of atomistic data [9]. In order to illustrate the thermal conductivity κ that results from MD simulations, we compute κ in this section as the opposite of the short-term averaged flux divided by the short-term averaged temperature gradient. This exercise results in a significant increase in the noise level since the value of temperature gradient samples could be close to zero. In macroscale heat transfer, thermal conductivity is known to be a strictly positive property. However, when κ is computed on the atomistic scale and plotted as a function of temperature, it is not surprising to observe that it takes negative values. Unlike at the macroscale scale, at such small atomistic time scales there is a substantial probability that the heat flux is in the same direction as the temperature gradient. The negative values decrease with increasing time averaging window sizes, as expected, and the probability density function (PDF) of κ is seen to approach a log-normal distribution.

3. MATHEMATICAL MODEL FORMULATION

In this section we describe the mathematical formulation for the inferred conductivity with specific implementation details. Major steps of the inference process are illustrated in Fig. 4.

3.1 Building the Heat Conduction Constitutive Law

For given L_a, T_{HOT}, and T_{COLD}, we gather short-time averaged samples of the thermal entities q, ∇T, and T at the mesh points after the simulation reaches a statistical steady state. As such, we obtain N_d independent noisy samples $\{q_j, \nabla T_j, T_j\}_{j=1}^{N_d}$. We use these samples as data in a Bayesian inference process to determine the posterior distribution of the parameters that relate q, ∇T, and T.

Given the assumption of a Fourier conduction law with temperature-dependent conductivity, Fig. 3 (thick black line) suggests a decreasing trend for κ with temperature. Thus, we postulate the following linear relationship:

$$\kappa = A - BT \tag{3}$$

where A and B and their associated uncertainties are to be determined by the Bayesian inference machinery. The heat flux q together with its noise have nonlinear dependence on T and ∇T as suggested by Eq. (3). Yet, we found that the noise level in all q, T, and ∇T decreases with short-term averaging consistent with the central limit theorem (CLT). Hence, we use a Gaussian noise model that relates A and B to the short-time averaged samples $\{q_j, \nabla T_j, T_j\}$ such that

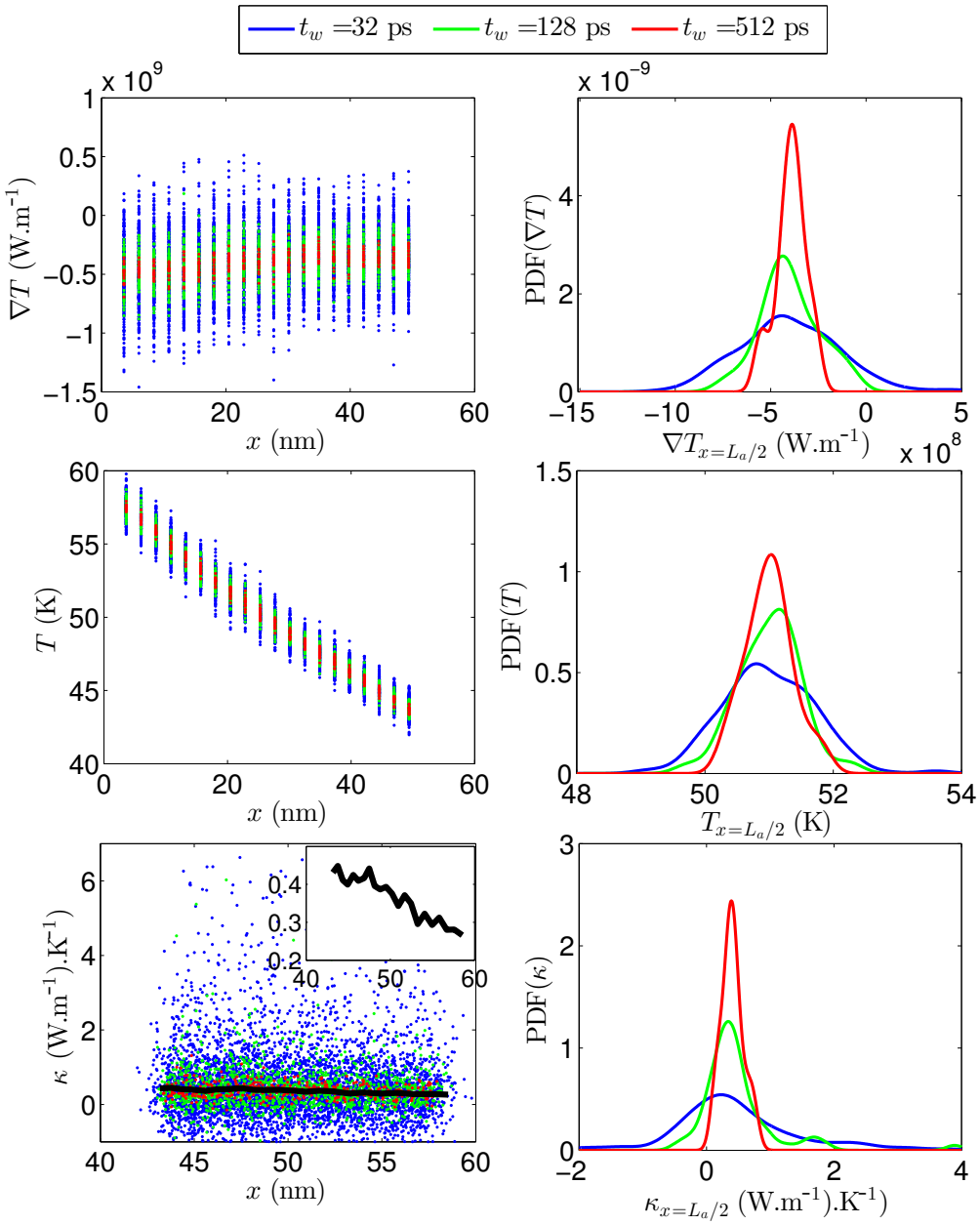

FIG. 3: Plots showing (left column): (top) temperature gradient field ∇T in the quasi-1D bar, (middle) temperature field T in the quasi-1D bar, (bottom) thermal conductivity κ as function of temperature, (right column): the PDFs of ∇T, T, and κ at $x = L_a/2$. Results are generated from the MD simulation for $T_{HOT} = 60$ K, $T_{COLD} = 40$ K, and $L_a = 53$ nm when the statistical steady state is reached for different time averaging window widths t_w, as indicated. The thick black line denotes the calculated mean thermal conductivity as a function of temperature.

$$q_j = -(A - BT_j)\nabla T_j + \sigma\eta_j = -A\nabla T_j + BT_j\nabla T_j + \sigma\eta_j \tag{4}$$

where the η_j represent the noise in the short-term MD sample averages. We do not consider any error due to model inadequacy since we assume that $-(A - BT_j)\nabla T_j$ is a good approximation to the true q_j's up to a zero-mean error

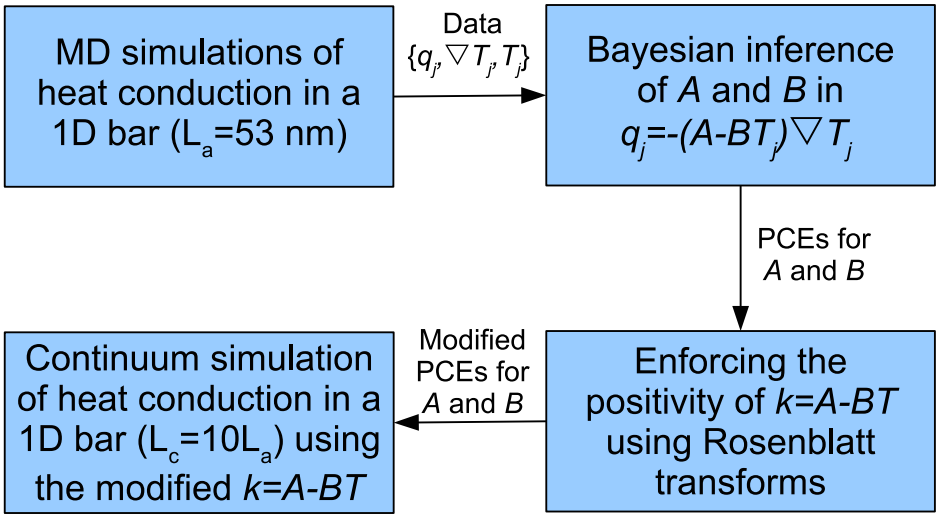

FIG. 4: Schematic illustrating the major steps of the constitutive law inference process.

term [17]. If for example the variation of the q_j's followed a more complicated trend as a function of the T_j's, a model inadequacy term would have been necessary to obtain an accurate inference. We neglect the correlation between the η_j's and assume they are independent and identically distributed standard normals. The variance σ^2 of the noise discrepancy is related to the fluctuations in the flux which depend on the temperature. σ^2 is independent of the number of samples N_d but decreases in a manner inversely proportional to the time averaging window t_w. We assume that the dependence of σ^2 on the temperature is weak for the temperature ranges considered in this paper, thus we take σ^2 to be a constant and infer it as a hyperparameter [29, 36] along with A and B. Bayes' rule is then written as

$$\mathcal{P}(A, B, \sigma^2|\boldsymbol{q}, \underline{d}) \propto \mathcal{P}(\boldsymbol{q}, \underline{d}|A, B, \sigma^2)\mathcal{P}(A, B, \sigma^2) \tag{5}$$

where $\boldsymbol{q} \in \mathbb{R}^{N_d}$ is the heat flux data q_j and $\underline{d} \in \mathbb{R}^{N_d \times 2}$ is a matrix containing the temperature and temperature gradient data obtained from the MD simulations such that $\underline{d}_j = (-\nabla T_j, T_j \nabla T_j)$ (the reader is referred to [16, 28, 29] for the details regarding the Bayesian inference methods used in this work).

The likelihood function based on Eq. (4) is written as:

$$\mathcal{P}(\boldsymbol{q}, \underline{d}|A, B, \sigma^2) = (2\pi\sigma^2)^{-N_d/2} \exp\left(-\frac{\boldsymbol{\epsilon}^T\boldsymbol{\epsilon}}{\sigma^2}\right) \tag{6}$$

where the elements of the discrepancy vector $\boldsymbol{\epsilon}$ are given by

$$\epsilon_j = \sigma\eta_j = q_j + (A - BT_j)\nabla T_j \tag{7}$$

We assume independent priors for A, B, and σ^2, i.e., $\mathcal{P}(A, B, \sigma^2) = \mathcal{P}(A)\mathcal{P}(B)\mathcal{P}(\sigma^2)$ and assign to both A and B an improper uniform prior [28] on $[-\infty, +\infty]$. Other choices of priors on A and B are possible; for example, we can choose a prior that handles the issue of the negative conductivities (see Fig. 3). However, with such priors, an analytical solution for the posterior is not available as discussed later in this section. For the hyperparameter, σ^2, leveraging the fact that there is a complete *a priori* ignorance about its value except that it cannot be negative, we assume a Jeffreys prior

$$\mathcal{P}(\sigma^2) = \frac{1}{\sigma^2} \tag{8}$$

If the likelihood (6) incorporates a large amount of data then the joint prior has a minimal role in the resulting posterior. Conversely, if the likelihood function only brings a small amount of data, the distribution of the resulting posterior is comparable to the prior. The effect of the amount of data will be discussed in Section 4.

Based on the chosen priors and the Gaussian noise model, the corresponding posterior over $\{A, B\}$ is a Student-t distribution [26, 37, 38] $\mathcal{S}(\gamma, \boldsymbol{\mu}, \underline{\mathsf{S}})$ where the mean vector $\boldsymbol{\mu}$, the number of degrees of freedom γ, and the scale matrix $\underline{\mathsf{S}}$ are given by [26]

$$\begin{aligned} \boldsymbol{\mu} &= \underline{\nu d}\boldsymbol{q} \\ \gamma &= N_d - 2 \\ \underline{\mathsf{S}} &= \frac{1}{\gamma}\left(\boldsymbol{q}^T\boldsymbol{q} - \boldsymbol{q}^T\underline{d}\boldsymbol{\mu}\right)\underline{\nu} \\ \underline{\nu} &= \left(\underline{d}^T\underline{d}\right)^{-1} \end{aligned} \tag{9}$$

Thus we can express A and B as

$$\begin{pmatrix} A \\ B \end{pmatrix} = \boldsymbol{\mu} + \underline{\Lambda}\boldsymbol{\xi} \sim \mathcal{S}(\gamma, \boldsymbol{\mu}, \underline{\mathsf{S}}) \tag{10}$$

where $\underline{\Lambda} \in \mathbb{R}^{2\times 2}$ is the lower-triangular Cholesky factor of the scale matrix $\underline{\mathsf{S}} = \underline{\Lambda\Lambda}^T$, and the vector $\boldsymbol{\xi} \in \mathbb{R}^2$ comprises independent and identically distributed (i.i.d.) random variables distributed according to $\mathcal{S}(\gamma, 0, 1)$. For large values of N_d, the Student-t distribution $\mathcal{S}(\gamma, 0, 1)$ approaches the normal distribution $\mathcal{N}(0, 1)$, such that, in this limit, $\mathcal{P}(A, B|\boldsymbol{q}, \underline{d})$ is close to a binormal distribution. A and B are finally written as first order ($p = 1$) PCEs [22–25, 39–42] with two stochastic dimensions ($n = 2$) representing the finite sampling noise [26]:

$$\begin{aligned} A &= \mu_1 + L_{11}\xi_1 \\ B &= \mu_2 + L_{21}\xi_1 + L_{22}\xi_2 \end{aligned} \tag{11}$$

where ξ_1 and ξ_2 follow $\mathcal{N}(0, 1)$ for large N_d [26]. The thermal conductivity $\kappa = A - BT$ thus becomes an uncertain variable and captures the uncertainty due to the limited amount of noisy MD simulations data. The uncertainty magnitude increases with the noise amplitude σ and decreases with N_d. In other words, the uncertainty magnitude increases with the noise amplitude σ and decreases with N_d. Identifying and modeling the intrinsic uncertainty is beyond the scope of this paper. Our main goal is to propagate this uncertainty in continuum simulation. However, since both A and B follow a Student-t distribution, they can take any value in $[-\infty, +\infty]$ allowing for κ to have negative values (see Fig. 3, bottom left) that are not mathematically feasible on the continuum level. Thus, additional operations have to be performed on the distributions of A and B to enforce the positivity of κ.

Remark 1. *The analytical solution obtained in this derivation is only feasible for a constant noise model in Eq. (4). If the noise level is assumed to have a dependence on the temperature and temperature gradient, a sampling method such as Markov Chain Monte Carlo should be used in order to construct the posterior in Eq. (5).*

3.1.1 Enforcing the Positivity of the Thermal Conductivity

The mathematics of the diffusion equation require a positive thermal conductivity everywhere. However, the procedure described above would result in negative values as, due to nanoscale fluctuations, the heat flux can follow the temperature gradient with nonzero probability. In order to guarantee that well-posed solutions exist for the continuum heat transport equations, we propose a set of operations on the distributions of A and B that allow enforcing the positivity of $\kappa = A - BT$ for a given range $T_{COLD} \leq T \leq T_{HOT}$. The effect of the procedures is to build new distributions for A and B. We proceed as follows:

1. We draw a large number, N_s, of realizations $(A, B)_i$ from the Student-t distributions of A and B derived in Section 3.1.
2. We eliminate from this set of realizations all the $(A, B)_i$ that satisfy $A_i - B_i T_{COLD} < 0$ or $A_i + B_i T_{HOT} < 0$.
3. Using kernel density estimation (KDE) [26, 43], we build new distributions for A and B from the remaining realizations.

Remark 2. *Another method to enforce the positivity of $\kappa = A - BT$ is to assign an appropriate prior on A and B in the Bayesian inference procedure in Section 3.1. The prior expression would then be*

$$\begin{aligned} \mathcal{P}(A,B) &= 0, \quad \text{for} \quad A - BT_{COLD} < 0 \quad \textit{or} \quad A + BT_{HOT} < 0 \\ \mathcal{P}(A,B) &= \frac{1}{(A_{max} - A_{min})(B_{max} - B_{min})}, \quad \text{otherwise} \end{aligned} \tag{12}$$

where A_{max}, A_{min}, B_{max}, and B_{min} determine a known range of variation of A and B. However, this approach would not not allow an analytical derivation for the distributions of A and B. It requires sampling the posterior in Eq. (5) using the Markov Chain Monte Carlo method. The samples that do not satisfy the constraints of Eq. (12) are rejected from the chain which is equivalent to the elimination performed in step 2 above.

Remark 3. *The initial number of realizations N_s should be large enough such that after eliminating the $(A,B)_i$, the number of the remaining realizations is sufficient for statistical significance of the resulting distribution. In our computations, we needed 10000 samples to obtain an accurate distribution of κ.*

Remark 4. *Another approach to eliminate negative thermal conductivities is by applying bigger time averaging scales t_w on the raw MD data (see Fig. 3). The approach we adopted in this study is more general and applicable to complicated and expensive MD simulations where it is not feasible to use high values of t_w, or when continuum models are desired to operate at time scales below the limiting behavior.*

3.1.2 Spectral Projection of the Thermal Conductivity

The distributions of A and B obtained in Section 3.1.1 are not guaranteed to follow any of the traditional distributions such as normal, log-normal, uniform, etc., particularly for the smaller sample sizes. Thus, the representation of the random variables A and B as PCEs generally requires higher order polynomials. We first employ a Rosenblatt transform to map the jointly distributed random variables A and B (after enforcing the positivity of their corresponding κ) into two independent uniform random variables. We then apply an approximate inverse Rosenblatt to compute the PC coefficients of A and B at a given expansion order. These operations are further detailed in the work of Sargsyan et al. [44, 45]. In this work we performed a convergence study and found that a fourth-order expansion ($p = 4$, see Table 1 below) with two stochastic dimensions ($n = 2$) can accurately represent both A and B. Hence, after enforcing the positivity of κ, A and B are written as

$$\begin{aligned} A &= \sum_{k=0}^{P} A_k \Psi_k(\xi_1, \xi_2) \\ B &= \sum_{k=0}^{P} B_k \Psi_k(\xi_1, \xi_2) \end{aligned} \tag{13}$$

where the Ψ_k's are Hermite polynomials given in [25, 46]. The number of terms in this expansion in this work is $P = 14$.

Figure 5 (top row) shows PDFs of A and B before and after enforcing the positivity of $\kappa = A - BT$. A and B were sampled from Eqs. (11) and (13), respectively. The PDFs were then constructed using KDE. After enforcing the positivity, the mean of κ increases, which physically implies that more heat is transported in the material when the effects of the backward propagating phonons are eliminated. The effect of enforcing the positivity of κ on its PDF decreases when more data are used in the inference process. A larger amount of data and/or a longer time averaging window reduces the uncertainty in the inferred values and shifts the PDF of κ shifts away from the negative regions. Figure 5 (bottom row) shows the posterior distribution for A and B before and after enforcing the positivity of κ. Enforcing the positivity of κ eliminates a significant region from the posterior support yielding to sharp edges at its boundary. The expected sharp edge of the PDF of κ near zero is slightly smoothed due to the use of the KDE method to construct the PDF. The area of the eliminated region decreases relatively with increasing time averaging window consistent with the observations in Fig. 3.

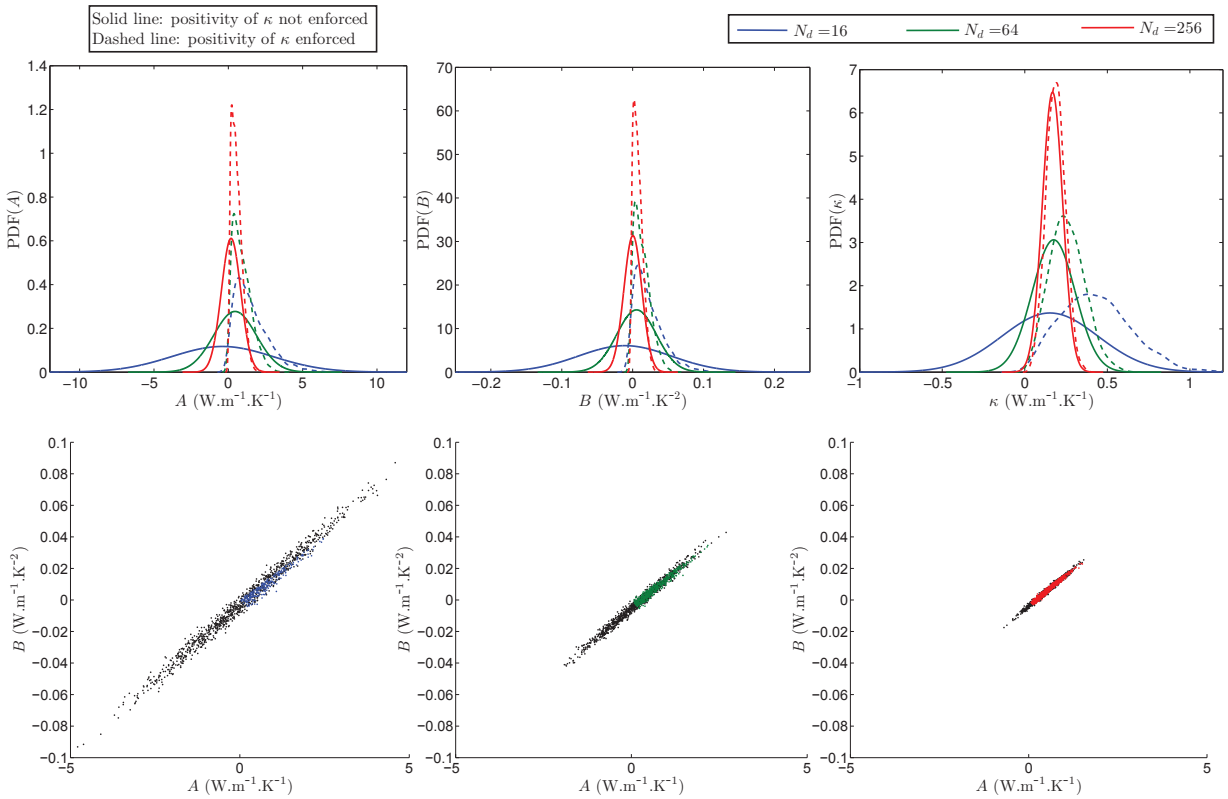

FIG. 5: Plots showing (top row) the PDFs of: (left) the inferred coefficient A, (middle) the inferred coefficient B, and the corresponding the thermal conductivity κ for $T = 50$ K. The PDFs of the coefficients are plotted before (solid line) and after (dashed line) enforcing the positivity of κ, and (bottom row) the posterior for A and B before (black dots) and after (color dots) enforcing the positivity of κ. Results are obtained from data averaged at $t_w = 8$ ps for different amounts N_d of data used for the inference, using four replica MD simulations for $T_{HOT} = 60$ K, $T_{COLD} = 40$ K, and $L_a = 53$ nm.

3.2 Numerical Implementation

We propose two sampling approaches to generate the Bayesian inference data $\{q_j, \nabla T_j, T_j\}_{j=1}^{N_d}$. We call the first approach "local" where the data is obtained from a single MD simulation with specified T_{HOT} and T_{COLD}. This was the case for the results generated in Fig. 5 where $T_{HOT} = 60$ K and $T_{COLD} = 40$ K. The other sampling approach called "global" relies on multiple MD simulations for the heat flux data. For convenience, we characterize each MD simulation by a mean temperature T_{MD} and a temperature difference ΔT_{MD} such that

$$\begin{aligned} T_{MD} &= \frac{T_{HOT} + T_{COLD}}{2} \\ \Delta T_{MD} &= T_{HOT} - T_{COLD} \end{aligned} \tag{14}$$

The local sampling approach used a single MD simulation with a fixed value of $\{\Delta T_{MD}, T_{MD}\}$ while the global sampling relies on 25 MD simulations covering the following parameter values: $\Delta T_{MD} = 10$, 20, 30, 40, and 50 K and $T_{MD} = 30$, 40, 50, 60, and 70 K, chosen just for illustration. Data obtained from these two approaches are plotted in Fig. 6. A higher ΔT_{MD} allows for better control of the heat flux during the MD simulation so there are fewer fluctuations in the resulting data. Moreover, the global sampling approach covers a broader range of both

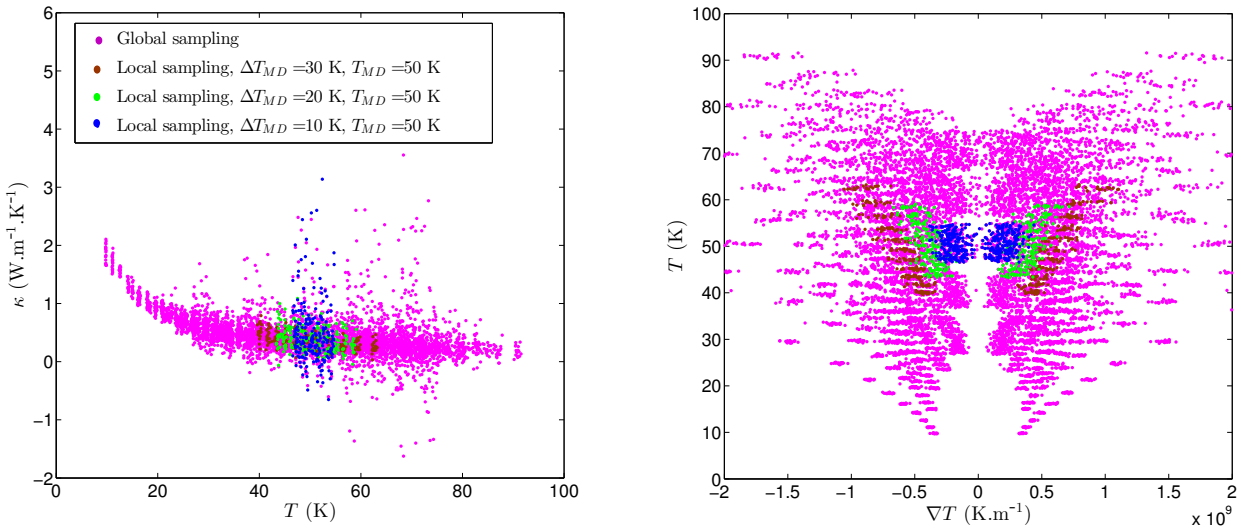

FIG. 6: Plots showing: (left) short-time averaged values of the thermal conductivity $\kappa_j = -q_j/\nabla T_j$ as a function of the temperature T_j, and (right) samples of the temperature T_j plotted as a function of the temperature gradient samples ∇T_j. Results are obtained from data averaged at $t_w = 512$ ps, using four replica MD simulations for different ranges of temperature by local and global sampling of the $\{\Delta T_{MD}, T_{MD}\}$ space, as indicated.

the temperature and temperature gradient. Figure 6 (left) shows that for the selected $\{\Delta T_{MD}, T_{MD}\}$ of the local sampling approach, the thermal conductivity κ decreases linearly with temperature, consistent with our assumption in Section 3. However, the global approach covers a larger temperature range and this linear trend is lost at lower temperatures beginning around $T = 20$ K. A linear dependence between κ and T was assumed for both "local" and "global" approaches in this paper but for the lower temperatures (e.g., $T = 20$ K), a more accurate prediction of the heat conduction law would then be obtained by assuming a higher order polynomial for the dependence of κ on T. We expect that in the case of a local sampling with high ΔT_{MD} and in the global sampling approach, a relatively smaller uncertainty will be obtained in the inferred A and B (the slope of the curve) with associated improved accuracy in the predicted heat conduction constitutive law for the form of the linear polynomial used in the present effort.

3.3 Propagating the Uncertain Constitutive Law into the Continuum Simulation

In this section we propagate the uncertainty in the thermal conductivity κ quantified in the previous section into a 1D continuum problem. Figure 7 shows the schematic of such a continuum simulation. The length scale of this simulation should be at least an order of magnitude bigger than the atomistic simulation length scale, i.e., $L_c \geq 10L_a$, to get into

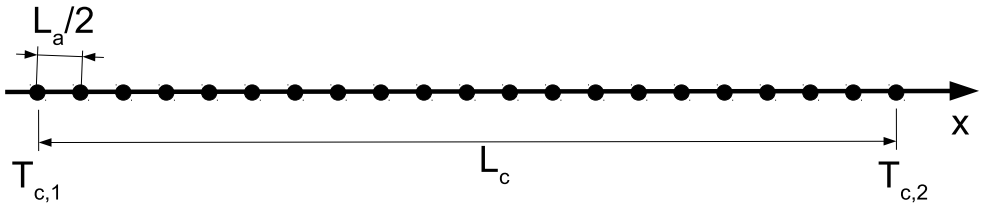

FIG. 7: Schematic showing the 1D continuum simulation domain. Dirichlet boundary conditions are imposed on the bar, as indicated. The bar has a length $L_c \geq 10L_a$. The black dots represent mesh points such that the mesh size is equal to $L_a/2$ (see Figs. 1 and 2).

a regime where the continuum formulation is valid. We discretize the continuum simulation domain such that the local shape functions spread over a length scale comparable to the atomistic simulation domain size, i.e., the mesh size h is given by $2h = L_a$. Consistent with the MD simulation, we impose Dirichlet boundary conditions on the continuum domain.

Given the assumed heat conduction constitutive law in Eq. (3), the partial differential equation (PDE) governing the heat conduction in the continuum scale 1D bar is written as:

$$\rho c_p \frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left[-(A - BT)\frac{\partial T}{\partial x}\right] \tag{15}$$

where the random variables A and B are given by the PCEs in Eq. (13). We propagate the uncertainty in A and B through the PDE above using a nonintrusive spectral projection (NISP) scheme (see [22–26] for details).

4. RESULTS

4.1 The PC Representation of the Constitutive Law

We first describe the PC coefficients of A and B, and the resulting thermal conductivity κ. The PCEs of A and B are given in Eq. (13). For a fixed temperature T, the PCE of κ is given by

$$\kappa = \sum_{k=0}^{P} (A_k - B_k T)\Psi_k(\xi_1, \xi_2) \tag{16}$$

We perform a convergence study on the moments of the distributions of A, B, and κ as a function of the expansion order p. The results of this study are given in Table 1 for the case with highest uncertainty considered in this study, in terms of mean and standard deviation, computed following [23–25]:

$$\begin{aligned} \mu_\kappa &= \kappa_0 \\ \sigma_\kappa &= \sqrt{\sum_{k=1}^{P} \kappa_k^2 \langle \Psi_k^2 \rangle} \end{aligned} \tag{17}$$

For $p = 4$, the values are essentially unchanged and thus we pick an expansion order $p = 4$ for the rest of the results shown in this section.

For brevity, we plot in Fig. 8 the first four PC coefficients of A as a function of the amount of data N_d for different time averaging windows t_w in MD simulations. The PC coefficients of B follow similar trends. As N_d and t_w increase, A_0, the coefficient representing the mean of A, approaches a constant value, indicating a more accurate solution. Concurrently, all other PC coefficients that describe the uncertainty decrease. Increasing t_w reduces the fluctuations in the instantaneous temperature, temperature gradient, and heat flux extracted from MD simulations. Similarly, increasing N_d improves the posterior prediction of the Bayesian inference procedure by increasing the likelihood of the data. The local sampling approach results in higher values of PC coefficients for all $k > 0$ than the global sampling approach. The latter covers a larger range of temperatures and temperature gradients, as discussed in Section 3.2, and results in more accurate predictions of the inferred A. In the local sampling approach, a smaller ΔT_{MD} results in higher PC coefficients for $k > 0$, i.e., higher uncertainty as expected from the observations on the

TABLE 1: The mean and standard deviation of A, B, and κ for different PC expansion orders

	$p = 2$	$p = 3$	$p = 4$	$p = 5$
(μ_A, σ_A) (W.m^{-1}.K^{-1})	(1.499,1.079)	(1.487,1.077)	(1.484,1.071)	(1.484,1.071)
(μ_B, σ_B) (W.m^{-1}.K^{-2})	(0.0222,0.0197)	(0.0216,0.0189)	(0.0212,0.0187)	(0.0212,0.0187)
$(\mu_\kappa, \sigma_\kappa)$ (W.m^{-1}.K^{-1})	(0.389,0.148)	(0.407,0.157)	(0.424,0.175)	(0.423,0.174)

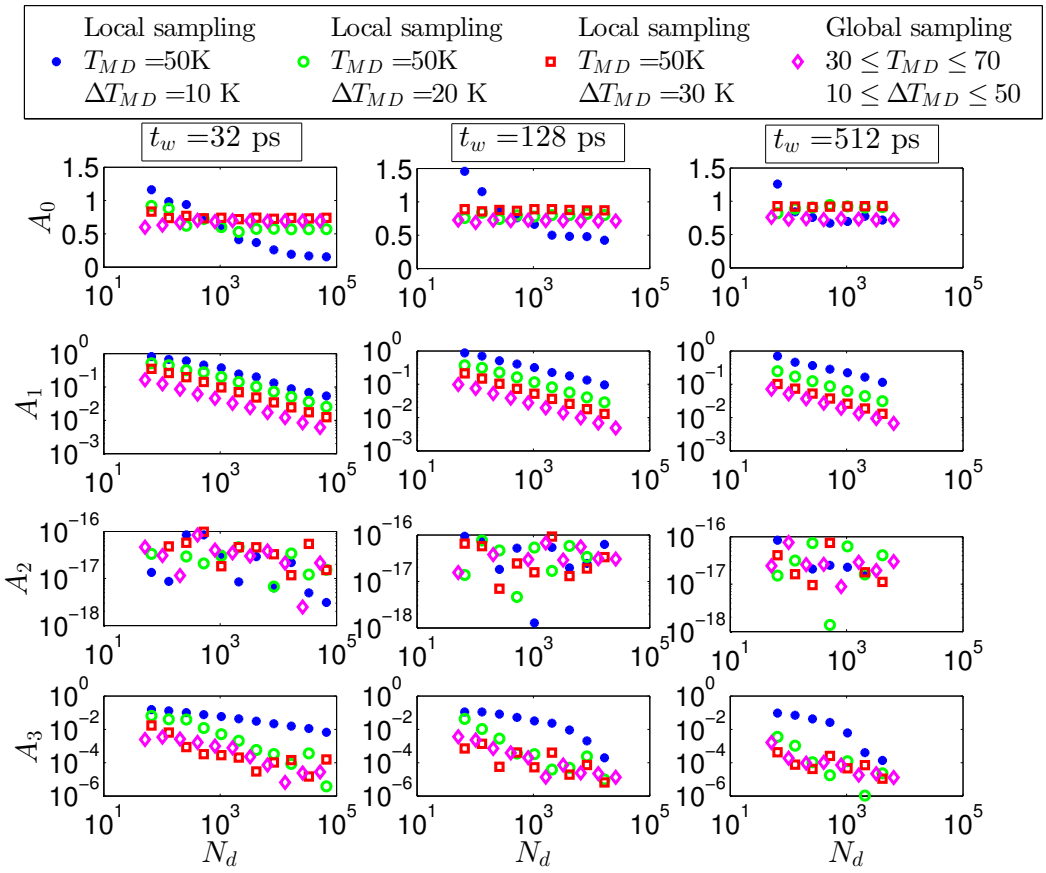

FIG. 8: Plots showing the first four PC coefficients as a function of the amount of data, N_d, used for the inference of the inferred coefficient A. Results are obtained after enforcing the positivity of κ from data averaged at different values of t_w, using four replica MD simulations for different ranges of temperature by local and global sampling of the $\{\Delta T_{MD}, T_{MD}\}$ space, as indicated.

data in Section 3.2. The logarithmic slope of the decrease in A_1 with N_d is very close to the value of 0.5 expected from the CLT. In fact, A_1 is closely related to the normal part of the distribution of A. However, the remaining PC coefficients exhibit a more complicated trend as a function of N_d. First, we notice that some modes, such as A_2, have negligible amplitudes, meaning they do not participate in assessing the distribution of A. Secondly, there is a rapid drop in A_3 occurring for $T_{MD} = 50$ K and $\Delta T_{MD} = 10$ K, at particular values of N_d and t_w. This drop in PC coefficients of A as a function of N_d is not seen when the positivity of κ is not enforced, i.e., if Rosenblatt tranforms are not performed. Hence, this drop is related to the changes in the PDFs of A shown in Fig. 5 after enforcing the positivity of κ.

We also describe the PC coefficients of $\kappa = A - BT$ and their dependence on N_d and t_w. The PC coefficients of κ are given in Eq. (16) where the temperature T is chosen to be $T = 50$ K. These PC coefficients are plotted in Fig. 9 before and after enforcing the positivity of κ. Note that in the former case, κ follows a Student-t process that we approximate by a Gaussian process as discussed in Section 3.1. Therefore, κ is described by a first-order PCE with two stochastic dimensions, i.e., it is described by three PC coefficients as shown in Fig. 9 (left column). Similarly to A_0, κ_0 approaches a constant value as N_d and t_w increase. κ_0 is seen to increase with t_w as it reaches a constant value. This increasing trend may be attributed to the fact that we are assuming a constant noise model in the inference of A, B, and κ (see the remark at the end of Section 3.1).

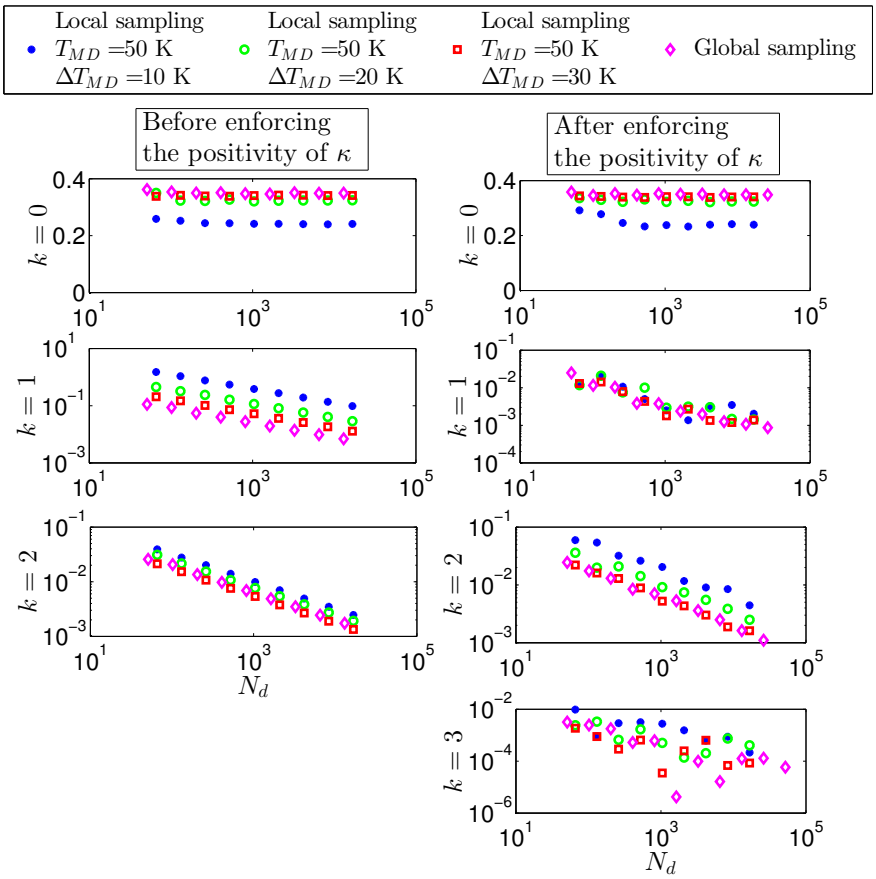

FIG. 9: Plots showing the PC coefficients as a function of the amount of data, N_d, used for the inference of the thermal conductivity κ for $T = 50$ K before (left) and after (right) enforcing its positivity. Only the first four PC coefficients are shown after enforcing the positivity of κ. Results are obtained from data averaged at $t_w = 128$ ps, using four replica MD simulations for different ranges of temperature by local and global sampling of the $\{\Delta T_{MD}, T_{MD}\}$ space, as indicated.

Before enforcing the positivity of κ, κ_1 and κ_2 follow the expected decreasing trend that is consistent with the CLT, i.e., the decrease is proportional to $\sqrt{N_d}$. After enforcing the positivity of κ, the PC coefficients exhibit different trends. While κ_1 and κ_2 show a linear decreasing trend with N_d, the decreasing trend of the higher modes is complicated and is probably related to the nature of the PDF of κ shown in Fig. 5.

4.2 The Inferred Constitutive Law

Given the PCE of κ, we now can compute the heat flux q as a function of the temperature T and the temperature ∇T. For a fixed T and ∇T, the PCE of q can be computed as

$$q_k = -(A_k - B_k T)\nabla T \tag{18}$$

Thus, the expectation of q at a given temperature and temperature gradient is given by its first PC coefficient following Eq. (17),

$$E[q] = -(A_0 - B_0 T)\nabla T \tag{19}$$

while the standard deviation of q is given by:

$$\sigma[q] = \sqrt{\sum_{k=1}^{P} \left[(A_k - B_k T)\nabla T\right]^2 \langle \Psi_k^2 \rangle} \tag{20}$$

For representative ranges of T and ∇T, we plot in Fig. 10 the isosurfaces of $E[q]$ and $\sigma[q]$ surface with A and B inferred from MD simulations data at $T_{MD} = 50$ K and $\Delta T_{MD} = 20$ K for different values of t_w. To aid in visualization, the data have been plotted over a symmetric range about zero in the temperature gradient. Scanning from left to right demonstrates how the mean values change as a bigger time averaging window is used in the inference, commensurate with a reduction in the standard deviation. Also importantly, it is possible to see how the enforcement

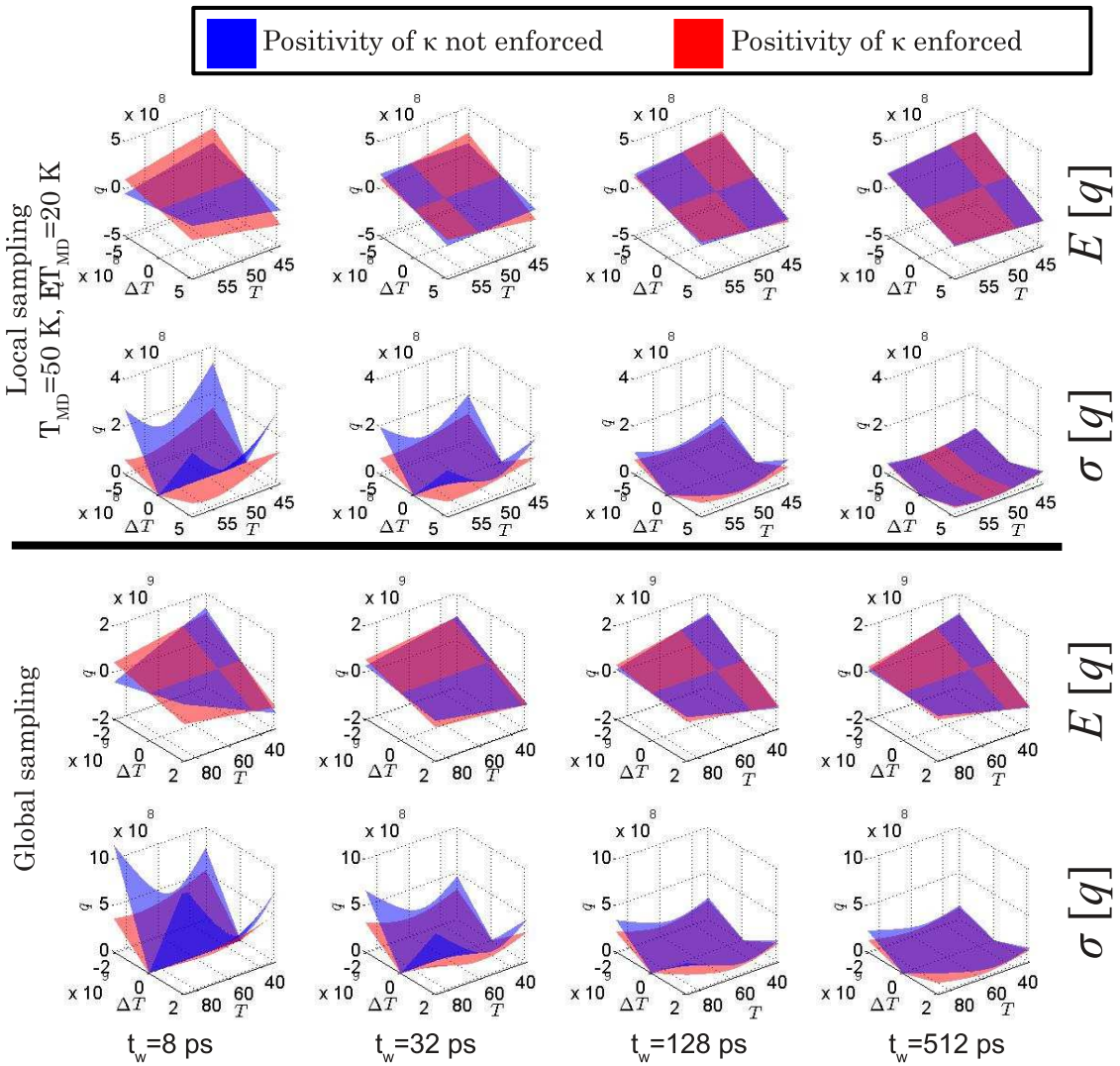

FIG. 10: Plots showing the expectation E and standard deviation σ of the flux $q = -(A - BT)\Delta T$ surface obtained after the inference of A and B, as a function of the amount time averaging window t_w. Results are obtained from $N_d = 16$ short-time averaged $(q_j, \nabla T_j, T_j)$ data, using four replica MD simulations by local and global sampling of the $\{\Delta T_{MD}, T_{MD}\}$ space, as indicated. q is given in W.m^{-2}, ∇T in K.m^{-1}, and T in K.

of the positivity of κ impacts the overall distribution. The effects of this enforcement are most clearly visible for smaller values of t_w because of the greater probability of phonons propagating with the temperature gradient rather than against it. Hence, when the conductivity is inferred from small spatial and/or temporal windows, the data-based restriction of κ to be positive effectively removes the negative conductivity samples and results in an overprediction of the heat flux. The results of all the "local" sampling approach ($\Delta T_{MD} = 20$ K) and the "global" approach are plotted together in Fig. 11, which shows similar standard deviation, while the globally sampled inference demonstrates a higher heat flux at lower temperatures. This is expected behavior for solids in which the phonon population increases at higher temperature, and the increased collision frequency reduces the conductivity.

4.3 Discussion

The generation of the MD data in the global sampling approach considered in this study costs 25 times more than the local sampling approach since more MD simulations had to be performed. This additional cost incurred benefits to the inference of the thermal conductivity. As mentioned in Section 3.2, the heat flux surface inferred using a global sampling approach occupies a bigger area in the $\{T, \nabla T\}$ space [see Fig. 6 (right)] allowing one to solve heat transfer problems with a variety of temperature ranges. The global sampling appraoch also resulted in a lower uncertainty level in the inferred parameters. The choice between a local and a global sampling depends on the scope of the continuum scale problem to be addressed, e.g., the temperature ranges and acceptable levels of uncertainty. There is a trade-off between spending an additional computational cost to generate data in a global sampling setting and affording restricted ranges of temperature with increased levels of uncertainty.

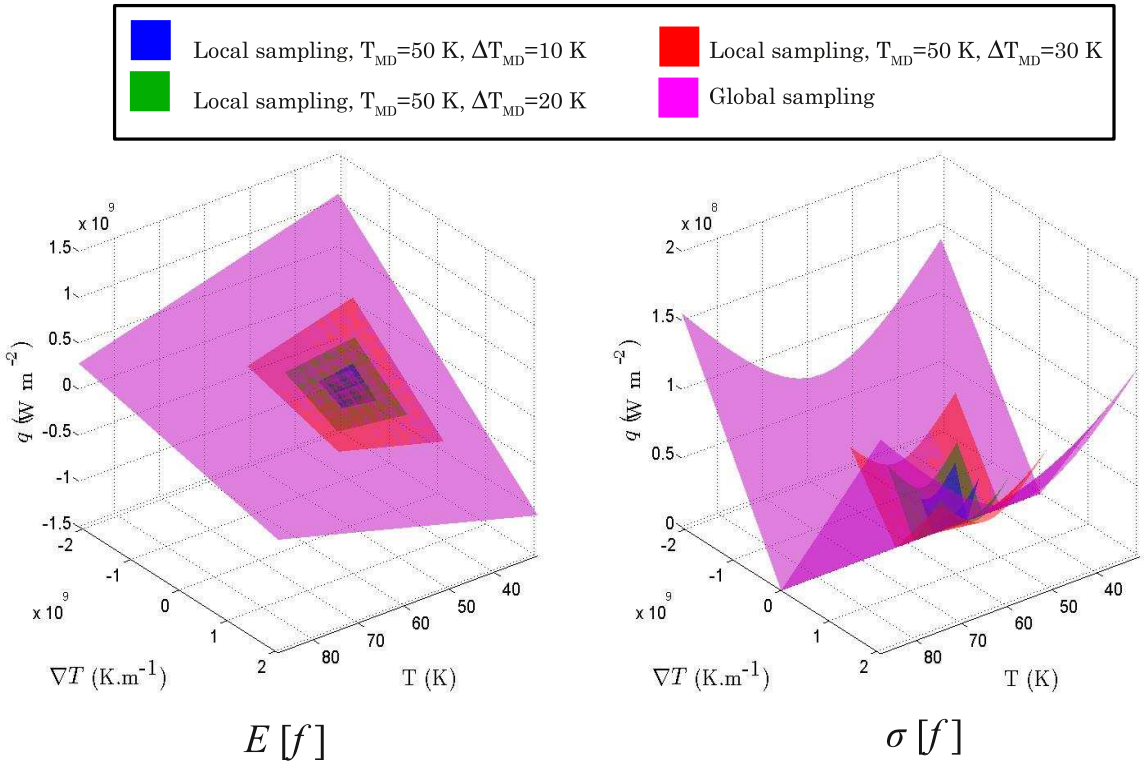

FIG. 11: Plots showing the expectation E and standard deviation σ of the flux $q = -(A - BT)\Delta T$ surface obtained after the inference of A and B. Results are obtained from $N_d = 16$ short-time $(f_j, \nabla T_j, T_j)$ data averaged at t_w =512 ps, using four replica MD simulations for different ranges of temperature by local and global sampling of the $\{\Delta T_{MD}, T_{MD}\}$ space, as indicated.

4.4 Continuum Simulation using the Inferred Constitutive Law

To demonstrate the model in an application, we propagate the uncertainty obtained in the thermal conductivity in a continuum scale simulation of a 1D bar as described in Section 3.3. We select the 1D bar length to be 10 times the one used in the atomistic simulation such that $L_c = 10L_a = 0.53$ μm. For given T_{MD}, ΔT_{MD}, time averaging window t_w, and number of samples N_d, we interpolate the values of the PC coefficients of A and B ($\kappa = A - BT$) from the curves plotted in Fig. 8. During the simulation, we assign Dirichlet boundary conditions $T_{c,1} = T_{HOT}$ and $T_{c,N} = T_{COLD}$ that are held fixed throughout the simulation, where T_{HOT} and T_{COLD} are calculated by Eq. (15) for the given T_{MD} and ΔT_{MD}.

Propagating the uncertainty κ in Eq. (15) results in an uncertain temperature field that we express by a PCE:

$$T(x,t) = \sum_{k=0}^{P} T_k(x,t)\Psi(\xi_1, \xi_2) \tag{21}$$

Based on Gauss-Hermite quadrature points [25], we calculate deterministic values of A and B that we use to compute a deterministic temperature field by solving the PDE [Eq. (15)]. We then use the ensemble of solutions to determine the PC coefficients $T_k(x,t)$ based of the non-intrusive spectral projection (NISP) approach [25]. The number of quadrature points per stochastic dimension should be at least equal to $n = (2p+1)/2$ [25] in order to obtain an accurate projection of the random entities. We thus choose $n = 5$ Hermite quadrature points per dimension, which implies 25 deterministic solutions of Eq. (15) are required. The values of ρ and c_p for solid Argon were calculated from the literature [34, 47] at $T = (T_{c,1} + T_{c,N})/2$.

The PCE of the resulting temperature field is computed at each time step until a steady state is reached. We report the results in terms of the mean temperature given by the first PC coefficient $E[T] = T_0$ and the standard deviation of the temperature computed as

$$\sigma[T] = \sqrt{\sum_{k=1}^{P} T_k^2 \langle \Psi^2 \rangle} \tag{22}$$

This resulting uncertainty has the same nature as the uncertainty in the inferred A and B in Section 3.1. It depends on the time averaging window t_w as well as the amount of data N_d used in the inference. Hence, it reflects the lack of knowledge in the variables due to the finite amount of data.

In order to evaluate the model, we also simulated the same continuum scale heat transfer in the 1D bar directly using MD. Here we extract the variables q, ∇T, and T using a localization function characterized with a length scale $\lambda_c = 10\lambda_a$ (see Section 2). Thus, we expect the fluctuation amplitude in the variables to be reduced by a factor of $\sqrt{10}$. This MD simulation requires substantial computational resources compared to the MD simulations described in Section 2 due to the increased system size $L_c = 10L_a$. Specifically, while the validation MD simulation required about 9.5 days on 256 computing nodes, all the MD samples together with the inference of the positivity enforced thermal conductivity and the continuum scale simulation needed about only 10 h using 256 nodes. We denote the large validation simulation by the "true" MD simulation. We first compare the results of these two simulations approaches of the large 1D bar in Fig. 12 in terms of the mean of the temperature. There is reasonable agreement between the mean of the temperature computed using the two approaches when a time averaging window $t_w > 128$ ps is used. Here, "reasonable" denotes that the large MD calculation is within the estimated uncertainty of its continuum approximations. At such large time scales, the mean of the thermal conductivity described by κ_0 reaches a constant value as depicted in Fig. 9. However, at smaller time scales, this mean is not accurately predicted and results in discrepancies between the modeled continuum and true MD solutions. Another undesirable property of this method is that the estimates converge monotonically from below as N_d and/or t_w increase, indicating that the estimate of the true solution is biased when using this method.

5. CONCLUSION

The goal of this paper was to assess the feasibility of determining unknown constitutive relationships from atomistic simulation data for use in continuum nanoscale models. As a test case, we chose one-dimensional heat conduction

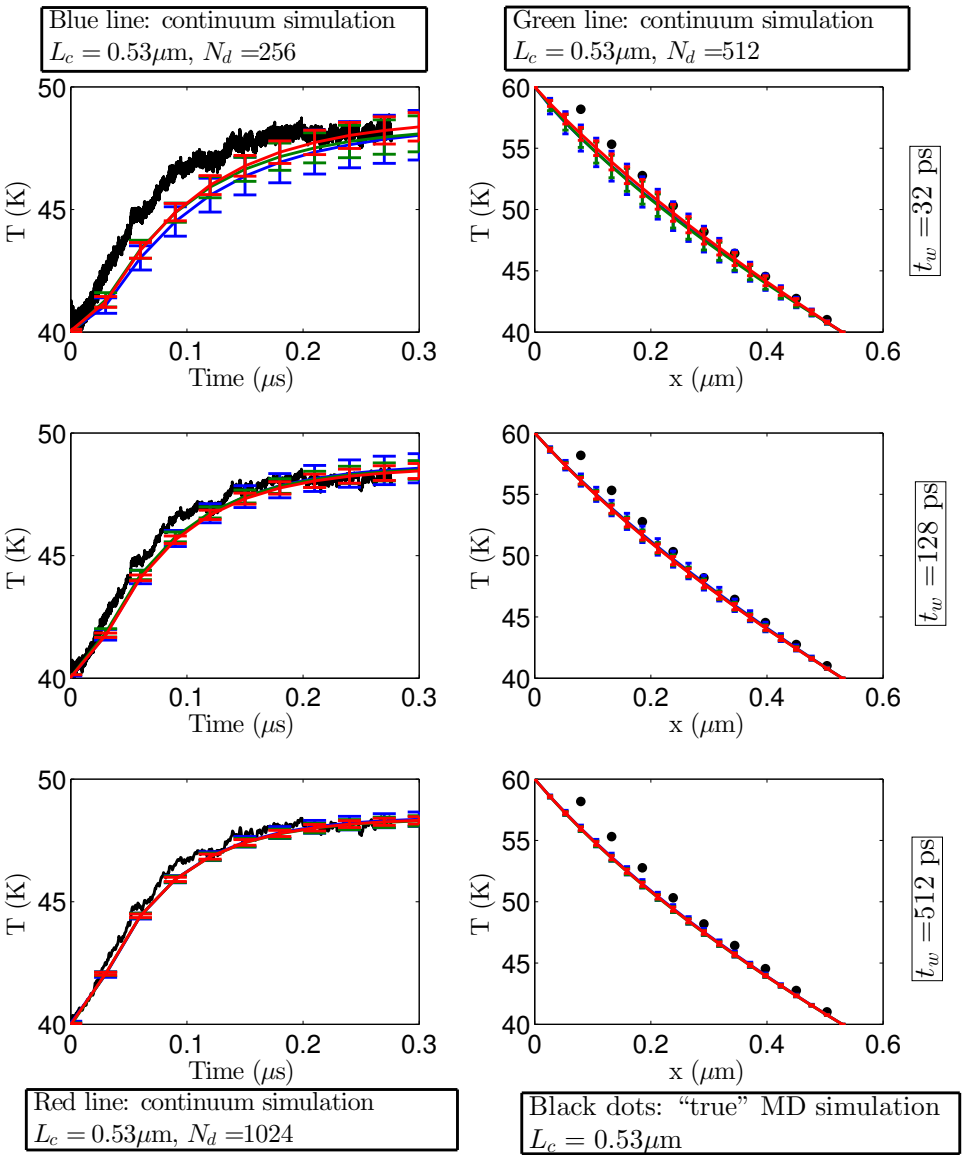

FIG. 12: Plots showing the (left column) temporal evolution of the mean and standard deviation of the temperature in the middle of a continuum scale 1D bar ($x = 0.26$ μm), and (right column) the steady-state mean and standard deviation of temperature distribution in the bar. Results are generated using two approaches: (blue, green, and red lines) by propagating the uncertain thermal conductivity inferred from MD simulations into a continuum simulation, and (black dots) the "true" MD simulation of the continuum scale bar ($L_c = 0.53$ μm). The error bars represent the calculated standard deviation in the temperature. Results are obtained for $T_{c,1} = 60$ K and $T_{c,N} = 40$ K and for different time averaging windows t_w, as indicated. The thermal conductivity κ is extracted from MD simulations for different values of N_d, as indicated.

as a model system because previous multiscale modeling efforts had established the connection between atomistic simulation and continuum Fourier thermal transport. Several approaches were reported. First, individual small MD simulations were performed and averaged to obtain estimates of the conductivity matching an assumed functional form. Efficient estimation was obtained by regarding the parameters defining the thermal conductivity as uncertain

variables represented by PCEs. These parameters were determined using Bayesian inference. Only a small number of MD replica simulations were required to obtain a sufficient number of samples by averaging the data over different time windows using the local approach. A second sampling strategy was also considered in which multiple MD samples were performed with different temperatures and temperature gradients to explore the global behavior of the conductivity. For problems in which the temperature range is known, the local approach is more efficient, while the global approach is a more robust choice when temperature distributions are not known *a priori*.

Direct inference of the thermal conductivity demonstrated that it is well approximated by a normal distribution with decreasing variance as the time averaging window size increased, demonstrating convergence to well-established classical theories. However, microscale heat transport is mediated by distributions of phonons which both along travel and against the temperature gradient. Because diffusion equations are ill-posed when the conductivity is negative, it is necessary to enforce the constraint that the conductivity is always positive. Doing so artificially enhanced the conductivity for smaller sample sizes and time averaging windows because negative heat fluxes were effectively removed, biasing the conductivity in the positive direction. It was shown that the process produced conductivity values which converged to their uncorrected values in the limits of large time windows. Enforcing positivity also increased the number of nonzero coefficients of the PC expansion since its order had to be slightly increased.

Assessment of the model was made by using the inferred conductivities in a 1D model of thermal transport from a hot temperature reservoir to a cooler one. This configuration enabled a direct comparison against the same physics modeled exclusively with molecular dynamics, which was taken to be the true solution. Importantly, the MD samples from which the conductivity was inferred were much smaller than the final system, demonstrating a significant potential cost savings.For larger sample sizes obtained either directly or through increased time averaging, the “true” MD simulation was generally within the uncertainty of the continuum scale simulation as estimated from the PCE. It was observed that as the MD averaging time window grew larger, the inferred conduction models converged to the mean solution, albeit in a biased fashion, and had correspondingly lower uncertainty.

Given these results, we conclude that sampling and inference methods have promise to be used in a wide variety of problems in which constitutive relations are needed to efficiently simulate a coarse-grained model at the nanoscale. In addition to high-fidelity measurements of the constitutive model, its uncertainties due to both finite sample sizes as well as physical stochasticity can be propagated through the solution process to assess confidence in the resulting solution within this framework. However, fluctuations at the nanoscale may render constitutive model forms suggested by either macroscale intuition or mean values from atomistic simulations inappropriate. Part 2 of this series will explore mechanisms to generalize the continuum model and the changes to the Bayesian processes necessary to exploit them.

ACKNOWLEDGMENTS

Sandia National Laboratories is a multiprogram laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the US Department of Energy’s National Nuclear Security Administration under contract DE-AC04-94AL85000. This work was supported by the Laboratory Directed Research and Development (LDRD), and its support is gratefully acknowledged. The authors would like to thank Dr. Jon Zimmerman for helpful discussions regarding the extraction of continuum properties from MD. The authors also thank Dr. Bert Debusschere for his constructive feedback on the paper, and Dr. Khachik Sargsyan for providing the software library used for the Rosenblatt transform.

REFERENCES

1. Ferreira, A. and Aphale, S., A survey of modeling and control techniques for micro- and nanoelectromechanical systems, *IEEE. Tran. Sys. Man. Cybernetics. Part C: Applications and Reviews*, 41(3):350–364, 2011.
2. Ricardez-Sandoval, L., Current challenges in the design and control of multiscale systems, *Can. J. Chem. Eng.*, 89(6):1324–1341, 2011.
3. Lan, J. and Li, G., A multiscale component mode synthesis approach for dynamic analysis of nanostructures, *J. Numer. Meth. Eng.*, 92(1):79–98, 2012.

4. Wang, X., Li, J., Lee, J., and Eskandarian, A., On the multiscale modeling of multiple physics, In: Li, S. and Ga, X. (Eds.), *Handbook of Micromechanics and Nanomechanics*, pp. 1–24, 2012.
5. Xionga, L., Denga, Q., Tuckerb, G., McDowellb, D., and Chena, Y., A concurrent scheme for passing dislocations from atomistic to continuum domains, *Acta. Mater.*, 60(3):899–913, 2012.
6. Irving, J. and Kirkwood, J., The statistical mechanical theory of transport processes. IV. The equations of hydrodynamics, *J. Chem. Phys.*, 18(6):817–829, 1950.
7. Hardy, R. J., Formulas for determining local properties in molecular-dynamics simulations: Shock waves, *J. Chem. Phys.*, 76(1):622–628, 1982.
8. Donev, A., Bell, J. B., Garcia, A. L., and Alder, B. J., A hybrid particle-continuum method for hydrodynamics of complex fluids, *SIAM Mult. Model. Simul.*, 8:871–911, 2010.
9. Zimmerman, J., Jones, R., and Templeton, J., A material frame approach for evaluating continuum variables in atomistic simulations, *J. Comput. Phys.*, 229:2364–2389, 2010.
10. Miller, R. and Tadmor, E., A unified framework and performance benchmark of fourteen multiscale atomistic/continuum coupling methods, *Modeling Simul. Mater. Sci. Eng.*, 17:053001, 2009.
11. Wagner, G., Jones, R., Templeton, J., and Parks, M., An atomistic-to-continuum coupling method for heat transfer in solids, *Comput. Methods Appl. Mech. Eng.*, 197:3351–3365, 2008.
12. Anciaux, G., Ramisetti, S., and Molinari, J., A finite temperature bridging domain method for md-fe copuling and application to a contact problem, *Comput. Methods Appl. Mech. Eng.*, 205:204–212, 2012.
13. Kotsalis, E., Walther, J., and Koumoutsakos, P., Control of density fluctuations in atomistic-continuum simulations of dense liquids, *Phys. Rev. E*, 76:016709, 2007.
14. Nie, X., Chen, S., and Robbins, M., Hybrid continuum-atomistic simulation of singular corner flow, *Phys. Fluids*, 16(10):3579–3591, 2004.
15. Donev, A., Bell, J., Garcia, A., and Alder, B., A hybrid particle-continuum method for hydrodynamics of complex fluids, *SIAM J. Mult. Model. Simul.*, 8(3):871–911, 2010.
16. Rizzi, F., Salloum, M., Marzouk, Y., Xu, R., Falk, M., Weihs, T., Fritz, G., and Knio, O., Bayesian inference of atomic diffusivity in a binary Ni/Al system based on molecular dynamics, *SIAM Mult. Model. Simul.*, 9:486–512, 2011.
17. Rizzi, F., Najm, H., Debusschere, B., Sargsyan, K., Salloum, M., Adalsteinsson, H., and Knio, O., Uncertainty quantification in MD simulations. Part I: Forward propagation, *SIAM Mult. Model. Simul.*, 10(4):1428–1459, 2011.
18. Rizzi, F., Jones, R., Debusschere, B., and Knio, O., Uncertainty quantification in MD simulations of concentration driven ionic flow through a silica nanopore. I. Sensivity to physical parameters of the pore, *J. Chem. Phys.*, 138:194104, 2013.
19. Rizzi, F., Jones, R., Debusschere, B., and Knio, O., Uncertainty quantification in MD simulations of concentration driven ionic flow through a silica nanopore. I. Ucertain potential parameters, *J. Chem. Phys.*, 138:194105, 2013.
20. Kim, M., Jang, H., Kim, H., and Cho, S., Multiscale adjoint design sensitivity analysis of atomistic-continuum dynamic systems using bridging scale decomposition, *Modeling. Simul. Mater. Sci. Eng.*, 21(3):035005, 2013.
21. Feng, E. and Crooks, G., Length of time's arrow, *Phys. Rev. Lett.*, 101:090602, 2008.
22. Ghanem, R. and Spanos, P., *Stochastic Finite Elements: A Spectral Approach*, Springer Verlag, 1991.
23. Le Maître, O., Knio, O., Najm, H., and Ghanem, R., A stochastic projection method for fluid flow. I. Basic formulation, *J. Comput. Phys.*, 173:481–511, 2001.
24. Le Maître, O., Reagan, M., Najm, H., Ghanem, R., and Knio, O., A stochastic projection method for fluid flow. II. Random process, *J. Comput. Phys.*, 181:9–44, 2002.
25. Le Maître, O. and Knio, O., *Spectral Methods for Uncertainty Quantification with Applications to Computational Fluid Dynamics*, Springer, Berlin, 2010.
26. Salloum, M., Sargsyan, K., Najm, H., Debusschere, B., Jones, R., and Adalsteinsson, H., A stochastic multiscale coupling scheme to account for sampling noise in atomistic-to-continuum simulations, *SIAM. Multi. Model. Simul.*, 10(2):550–584, 2011.
27. Gelman, A., Carlin, J., Stern, H., and Rubin, D., *Bayesian Data Analysis*, CRC/Chapman and Hall, 2003.
28. Sivia, D. and Skilling, J., *Data Analysis, A Bayesian Tutorial*, 2nd ed., Oxford Science, Oxford, 2006.

29. Marzouk, Y., Najm, H., and Rahna, L., Stochastic spectral methods for efficient bayesian solution of inverse problems, *J. Comp. Phys.*, 224:560–586, 2007.
30. Plimpton, S., Fast parallel algorithms for short-range molecular dynamics, *J. Comp. Phys.*, 117:1–19, 1995.
31. Allen, M. and Tidlesley, D., *Computer Simulation of Liquids*, Oxford Science Publications, Oxford, 1989.
32. Leach, A., *Molecular Modelling: Principles and Applications*, Pearson, Essex, 2001.
33. Frenkel, B. and Smit, B., *Understanding Molecular Simulation: From Algorithms to Applications*, Elsevier, 2002.
34. Wagner, G., Jones, R., Templeton, J., and Parks, M., An atomistic-to-continuum coupling method for heat transfer in solids, *Comput. Methods Appl. Mech. Eng.*, 197:3351–3365, 2008.
35. Templeton, J., Jones, R., and Wagner, G., Application of a field-based method to spatially varying thermal transport problems in molecular dynamics, *Modeling. Simul. Mater. Sci. Eng.*, 18:1–22, 2010.
36. Mohammad-Djafari, A., Bayesian inference for inverse problems, in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, Vol. 21, pp. 477–496, 2002.
37. Rowe, D., *Multivariate Bayesian Statistics: Models for Source Separation and Signal Unmixing*, Chapman-Hall, London, 2003.
38. Grimmett, G. and Stirzaker, D., *Probability and Random Processes*, Oxford, London, 2011.
39. Wiener, S., The homogeneous chaos, *J. Comput. Phys.*, 224:560–586, 2007.
40. Cameron, R. and Martin, W., The orthogonal development of nonlinear functionals in series of Fourier-Hermite functionals, *Ann. Math.*, 48:385–392, 1947.
41. Janson, S., *Gaussian Hilbert Spaces*, Cambridge University Press, Cambridge, 1997.
42. Abramowitz, M. and Stegun, I., *Handbook of Mathematical Functions*, Dover, New York, 1970.
43. Bowman, A. and Azzalini, A., *Applied Smoothing Techniques for Data Analysis*, Oxford University Press, New York, 1997.
44. Sargsyan, K., Debusschere, B., Najm, H., and Le Maître, O., Spectral representation and reduced order modeling of the dynamics of stochastic reaction networks via adaptive data partitioning, *SIAM. J. Sci. Comput.*, 31(6):4395–4421, 2010.
45. Sargsyan, K., Safta, C., Debusschere, B., and Najm, H., Uncertainty quantification given discontinuous model response and a limited number of model runs, *SIAM. J. Sci. Comput.*, 34(1):B44–B64, 2012.
46. Ghanem, R. and Spanos, P., *Stochastic Finite Elements: A Spectral Approach*, Springer Verlag, New York, 1991.
47. Dobbs, E. and Jones, G., Theory and properties of solid argon, *Rep. Prog. Phys.*, 20(1):516–564, 1957.