MULTISCALE MODEL. SIMUL.
Vol. 19, No. 1, pp. 533–549

NONLINEAR CONSTITUTIVE MODELS FOR NANO-SCALE HEAT CONDUCTION*

WEIQI CHU† AND XIANTAO LI‡

Abstract. We present a first-principle based approach that leads from a many-particle description to a nonlinear, stochastic constitutive relation for the modeling of transient heat conduction processes at the nano-mechanical scale. By enforcing statistical consistency, in that the statistics of local energy is consistent with that from an all-atom description, we identify the driving force as well as the model parameters in these generalized constitutive models. The connections to established generalized constitutive relations, including Cattaneo–Vernotte-type models, will be demonstrated.

Key words. generalized heat conduction, Mori–Zwanzig formalism, nano-scale modeling

AMS subject classifications. 60H10, 65C20

DOI. 10.1137/19M1257664

1. Introduction.

Heat conduction is a fundamental process that occurs in most mechanical and biological systems. Thermal properties in nano-mechanical systems have particularly significant impacts on the performances of nanodevices. However, on such scales, the transport of thermal energy exhibits a wide variety of phenomena that are different from macroscopic observations, as documented in many recent publications. Examples include, but are not limited to, the size dependence of thermal conductivity, heat pulse propagation, and delay phenomena [3, 46, 1, 7]. In particular, there are overwhelming experimental observations that indicate the breakdown of the conventional model of heat conduction [34, 3].

At the phenomenological level, a remarkable advancement in modeling non-Fourier behavior is the Cattaneo–Vernotte (CV) model [6, 63], which eliminates the paradox of the infinite speed of temperature propagation. Further generalizations, e.g., the Guyer–Krumhansl (GK) model [23], the Tzou model [62], and the extended thermodynamics models [1], all involve auxiliary equations for heat flux. Other extensions include nonlinear heat conduction models, where heat conductivity is temperature-dependent, which may give rise to traveling wave solutions [51], and stochastic heat equations, which have been studied extensively in the area of stochastic PDEs [52, 65].

The availability of molecular dynamics (MD) models has encouraged a great deal of effort to simulate and understand heat conduction problems directly at the nano-mechanical scale as an alternative to conventional methods. MD simulations can be done either at equilibrium, from which the heat conductivity is estimated from the Green–Kubo formula [60], or under a nonequilibrium setting, where the temperature is controlled at the boundary, and upon reaching a steady state, the heat conductivity can be estimated from the average heat fluxes. The literature on the subject is abundant, and we only refer the reader to the studies [18, 8, 27, 31, 39, 38, 44, 45, 64]

*Received by the editors April 22, 2019; accepted for publication (in revised form) January 27, 2021; published electronically March 24, 2021.
https://doi.org/10.1137/19M1257664
Funding: This research was supported by the NSF under grants DMS-1522617 and DMS-1619661.
†Department of Mathematics, University of California, Los Angeles, Los Angeles, CA 90095 USA (chu@math.ucla.edu).
‡Department of Mathematics, Pennsylvania State University, University Park, PA 16802 USA (xli@math.psu.edu).

and the references therein.

Meanwhile, it is still not understood, even at the conceptual level, how the MD models, a many-particle and time-reversible description, would conspire to the aforementioned generalized heat conduction models. In addition to its theoretical significance, bridging the two descriptions is also of practical importance. For example, unlike the Green–Kubo formula [60] for heat conductivity, the microscopic definitions of the parameters in the CV, GK, or Tzou models have not been established.

This paper presents a first-principle based derivation of generalized, nonlinear, stochastic heat conduction models, aiming to understand the passage from the atomic level to a larger scale description. We try to identify the transient processes of local energies by solving these reduced heat conduction models, without implementing full MD simulations, recording trajectories of all particles and computing energies from velocities and displacements at each step. We formulate generalized constitutive relations for non-Fourier heat conduction processes, in the form of closed stochastic equations of energy or heat flux. We attempt to identify the motivation for the nonlinear form, seek the origin of the random noise, and make connections to the CV and GK models.

To derive such a stochastic model from full MD, we will employ the Mori–Zwanzig (MZ) projection formalism [48, 72], which has been quite successful in deriving coarse-grained models [9, 15, 26, 28, 30, 58, 21, 25, 71]. For example, a lot of effort has been made to construct efficient approximations of the memory integral [15, 25, 42, 53, 69, 70, 50]. For the current problem, to arrive at a nonlinear constitutive relation, we propose a new projection procedure, to be referred to as an *entropy-based* projection. The main departure from Mori's conventional projection is that the current projection procedure results in a nonlinear driving force in the reduced models with which modeling non-Gaussian statistics is more straightforward. An added advantage is that the new projection leads to a linear convolution for the memory integral, which makes our models easier to handle than the state-dependent form from Zwanzig's projection.

Furthermore, a systematic approximation, similar to Mori's continued-fraction approach [47], can be constructed to approximate the memory term, which subsequently leads to the generalized heat conduction models. Our current approximation eliminates the memory by introducing auxiliary variables in a self-consistent manner. The motivation is threefold: First, it drastically simplifies the calculation of the memory integral and can be viewed as a fast summation method. Second, this procedure enables a simple alternative to model the random noise in such a way that the marginal density is preserved. Finally, with such an approximation, the models can be compared directly to some of the existing generalized constitutive relations.

The last part of the paper examines the statistics associated with the stochastic models. We demonstrate that the stochastic models can reproduce the one- and two-point statistics of the local energy from the full MD model. This paper, however, will not address some other issues, including the existence of a traveling wave [51], the dependence of the heat conduction properties on the system size and geometry, and the rigorous continuum limits of the stochastic models [24].

The rest of the paper is organized as follows. Section 2 presents a new projection formalism that leads to a reduced description of the local energy. This constitutes the basis for modeling energy transport. In addition, by introducing local approximations, we derive a hierarchy of generalized heat conduction models that can be readily compared to existing models. In section 3, we validate the approach by examining the statistics from the reduced models.

2. The derivation of heat conduction models.

2.1. Numerical observations. We start with observations from some numerical experiments. The first example is a 1D Fermi–Pasta–Ulam (FPU) model, with potential given by

$$V = \sum_j \ell(x_{j+1} - x_j), \quad \ell(u) = \frac{1}{2}u^2 + \frac{1}{4}u^4, \tag{2.1}$$

which has been studied extensively in statistical physics, e.g., [40]. The second example is a single-wall carbon nanotube, with interactions modeled by the Tersoff potential [59]. In terms of heat conduction process, both systems can be viewed as one-dimensional or quasi one-dimensional chain structures, even though atoms in the nanotube system move in a three-dimensional space.

In both examples, the chain structures are divided into blocks geographically $\{\Omega_i\}$, as shown in Figure 1, in order to define local energies (denoted by $\boldsymbol{a}$) associated with them.

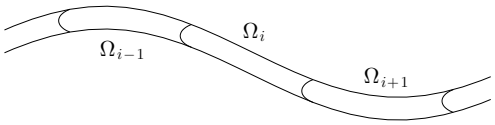

FIG. 1. *Diagram of a quasi-1D chain system with partition. Ω_i indicates the domain of the ith block.*

The ith component of $\boldsymbol{a}$ represents the energy associated with the ith block, consisting of the sum of potential energy and kinetic energy of atoms in the ith block. When the potential energy involves two atoms from two different blocks, the energy will be divided equally into the two blocks. For example, in the nearest-neighbor interaction models, the local energy is defined as

$$a_i = \sum_{j\in\Omega_i} \frac{1}{2}m|\boldsymbol{v}_j|^2 + \frac{1}{2}\ell(\boldsymbol{x}_{j+1} - \boldsymbol{x}_j) + \frac{1}{2}\ell(\boldsymbol{x}_j - \boldsymbol{x}_{j-1}), \tag{2.2}$$

where $\boldsymbol{x}_j, \boldsymbol{v}_j \in \mathbb{R}^d$ are the displacement and velocity of the jth atom and $\ell(\cdot)$ is the pair potential. In the FPU chain example $d = 1$ and in the nanotube example $d = 3$. The general scheme for energy partition, especially for multibody interactions, has been studied in [66].

We observe the statistics of local energies generated from a direct MD simulation under the canonical ensemble (NVT), $f(\boldsymbol{x}, \boldsymbol{v}) = e^{-\beta E(\boldsymbol{x},\boldsymbol{v})}/Z$, where $\beta = 1/k_BT$ is the inverse temperature and E is the energy of the whole system. Two ending blocks of the chain are connected with Nosé–Hoover thermostats [49] at the same temperature in order to force the whole system to arrive in an NVT ensemble. For the numerical methods, a sixth-order symplectic integrator is used to solve the Nosé–Hoover model, where the integrator is constructed based on a second-order operator-splitting method [61], followed by an extrapolation scheme [68]. After the system reaches a steady state, the thermostats are removed and periodic boundary conditions are imposed on both ends, i.e., the left-end block is connected to a block cell in the left, which behaves exactly the same as the right-end block, and vice versa. Unless otherwise mentioned, atomic units are used throughout the paper.

For the FPU chain, we consider a system of a total of 500 atoms with one block containing 10 atoms. The equilibrium temperature is set to be 1.5 a.u. For the nanotube example, the whole system has 1920 atoms with each block containing 16 atoms. The temperature is set to $300K$, which is converted to atomic units in the simulations. The displayed values of the energy follow the atomic units used in the Tersoff potential (eV).

We run MD simulations long enough to ensure that the system reaches a steady state and observe the probability density function (PDF) of the local energy a_i, as shown in Figure 2. The energy has been shifted to have a zero mean.

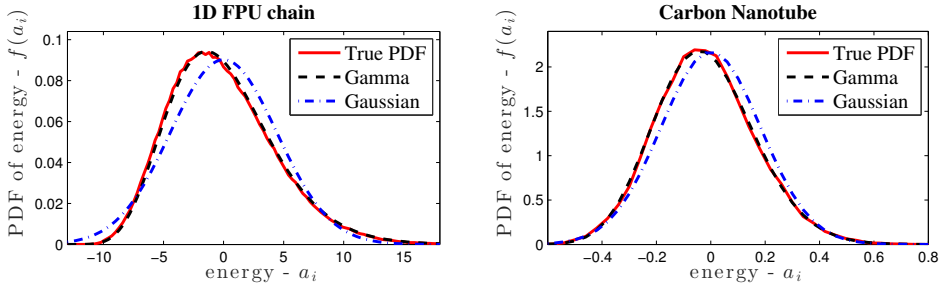

FIG. 2. *The PDFs of shifted local energy of the 5th block are compared with fitting Gaussian and Gamma distributions. We collect energy trajectories in a total of n_T steps from direct MD simulations, after reaching its equilibrium state. The stepsize is $\Delta t = 0.02$ a.u. The true PDFs are normalized histograms of the data, which fits into Gamma $f(x) = \frac{1}{Z}(x+\mu)^{\alpha-1}\exp\left(-\eta(x+\mu)\right), \mu = \alpha/\eta$, and Gaussian $f(x) = \frac{1}{Z}\exp\left(-x^2/2\sigma^2\right)$ distributions with correct mean and variance. The left plot shows the FPU chain example with a total number of steps $n_T = 10^7$. The right plot shows the nanotube example, with $n_T = 4 \times 10^6$. The Gamma distribution property is observed for the local energy in each block, which is expected due to the periodic property of the system. For the FPU chain example, fitting parameters are $\alpha = 9.8006, \eta = 0.7076$, and $\sigma = 4.4476$. For the carbon nanotube example, fitting parameters are $\alpha = 62.6716$, $\eta = 43.0849$, and $\sigma = 0.1837$.*

One interesting finding is that the statistics of the energy in such a regime is non-Gaussian. The PDF actually fits better to a shifted Gamma distribution,

$$\rho(a_i) \propto (a_i + \mu)^{\alpha-1} e^{-\eta(a_i+\mu)}. \tag{2.3}$$

It has also been observed, with two-dimensional histogram plots, that the local energy is almost independent among the blocks. A reasonable ansatz for the joint PDF is given by

$$\rho(\boldsymbol{a}) \propto \prod_{i=1}^{n_{\text{block}}} (a_i + \mu)^{\alpha-1} e^{-\eta(a_i+\mu)}. \tag{2.4}$$

Here (α, η) are the parameters in the Gamma-distribution which can be determined according to numerical data or empirical theories, and $\mu = \alpha/\eta$ represents the shift to ensure a zero mean of the energy. The independence in the local energy has also been observed for macromolecules [22].

For systems consisting of identical particles, as the block size increases, the local energies tend to exhibit a Gaussian property. Due to the central limit theorem, a non-Gaussian statistics emerges when the block size remains at the nano-mechanical scale. However, for complex systems with multiple types of particles, such as macromolecules, proteins, etc., non-Gaussian statistics has been observed at a larger scale [57].

In the following, we start from the MD setting and present a first-principle based method which is able to model the non-Gaussian processes and ensure the correct statistics at equilibrium.

2.2. A new Mori–Zwanzig projection formalism. The MZ formalism [48, 72] has recently re-emerged as a powerful tool to derive coarse-grained models based on a full atomistic description—the molecular dynamics model involving the position $\boldsymbol{x}(t)$ and the velocity $\boldsymbol{v}(t)$ of atoms:

$$\dot{\boldsymbol{x}} = \boldsymbol{v}, \quad m\dot{\boldsymbol{v}} = -\nabla V(\boldsymbol{x}), \tag{2.5}$$

Koopman

where $\boldsymbol{x}, \boldsymbol{v} \in \mathbb{R}^{3N}$ and N is the number of atoms. V is the total potential energy. Given the trajectories of all the atoms, $\big(\boldsymbol{x}(t), \boldsymbol{v}(t)\big)$, a preselected quantity of interest,

$$\boldsymbol{a}(t) := \boldsymbol{\phi}(\boldsymbol{x}(t), \boldsymbol{v}(t)), \tag{2.6}$$

serves as the coarse-grain (CG) variables whose dimension M is much less than the original dimension $6N$. The MZ formalism is designed to derive a reduced equation for $\boldsymbol{a}(t)$. The implicit dependence of $\boldsymbol{a}(t)$ on the initial state, $\big(\boldsymbol{x}_0, \boldsymbol{v}_0\big)$, will not be invoked unless necessary.

More specifically, the MZ formalism yields an exact equation,

$$\dot{\boldsymbol{a}}(t) = e^{t\mathcal{L}}\mathcal{P}\mathcal{L}\boldsymbol{a}(0) + \int_0^t e^{(t-\tau)\mathcal{L}}\mathcal{P}\mathcal{L}\boldsymbol{F}(\tau)d\tau + \boldsymbol{F}(t). \tag{2.7}$$

Here $\mathcal{L}$ is the Liouvillian, $\mathcal{P}$ is a projection operator, and $\mathcal{Q} = \mathcal{I} - \mathcal{P}$ is the complementary projection operator. $\boldsymbol{F}(t) = e^{t\mathcal{Q}\mathcal{L}}\mathcal{Q}\mathcal{L}\boldsymbol{a}(0)$ is typically regarded as a noise. The second term on the right-hand side indicates the history-dependence as a result of eliminating the excessive degrees of freedom. Interested readers are referred to [11, 9] for the mathematical derivations. What distinguishes the Mori and Zwanzig approaches is the choice of the projection operator $\mathcal{P}$ [9]. More specifically, Mori [48] formulated the problem in a Hilbert space and defined $\mathcal{P}$ as the projection onto the subspace spanned by $\boldsymbol{a}(0)$. With this projection, the first term on the right-hand side of (2.7) is a linear function of $\boldsymbol{a}(t)$, namely,

$$e^{t\mathcal{L}}\mathcal{P}\mathcal{L}\boldsymbol{a}(0) = \Omega\boldsymbol{a}(t). \tag{2.8}$$

In addition, the memory term can be written as a *linear* convolution,

$$\int_0^t e^{(t-\tau)\mathcal{L}}\mathcal{P}\mathcal{L}\boldsymbol{F}(\tau)d\tau = \int_0^t \theta(t-\tau)\boldsymbol{a}(\tau)d\tau, \tag{2.9}$$

which can be conveniently approximated by Mori's continued-fraction method [47], or a general rational function approximation. However, due to the linearity of the model, it is difficult to obtain non-Gaussian statistics, unless nontrivial approximations are introduced for the random noise [13]. Ciccotti and Ryckaert [14] proposed applying Mori's projection, but to a different quantity. The procedure yields a generalized Langevin equation with nonlinear force. However, the relation between the random noise and the memory kernel is quite complicated.

On the other hand, nonlinear models can be obtained from Zwanzig's approach [72], where the projection operator is defined as a conditional expectation that provides the optimal prediction in the L^2 norm. With this approach, the first term on

the right-hand side of (2.7) can often be related to the mean force associated with the CG variable $\boldsymbol{a}$ [15], which is regarded as an advantage from the perspective of ensuring correct statistics. However, the main difficulty in Zwanzig's formulation is the treatment of the memory term. In general, the integral cannot be written as a convolution and the continued-fraction approach breaks down. Short-time approximations, such as t-models [11, 12], as well as renormalization techniques to ensure long-time accuracy [53], have demonstrated some success for the reduction of some PDE models. Such approximations are difficult to implement for MD models, since the procedure involves high order derivatives of the potential energy, for which explicit forms are often not available. It is also not clear how the random noise can be treated within the framework of t-models.

In this paper, we propose another projection formalism in order to reconcile the two approaches. First, to ensure the consistency with the true statistics, we first write the PDF as

$$\rho(\boldsymbol{a}) = \frac{\exp[-W(\boldsymbol{a})]}{\int \exp[-W(\boldsymbol{a})] d\boldsymbol{a}}. \tag{2.10}$$

$W(\boldsymbol{a})$ is uniformly determined by the PDF up to a constant.

Based on our observation from the Gamma-distribution (2.4), we can define the corresponding $W(\boldsymbol{a})$ as

$$W(\boldsymbol{a}) = \sum_{i=1}^{n_{\text{block}}} w(a_i), \tag{2.11}$$

where

$$w(\xi) = \eta(\xi + \mu) - (\alpha - 1)\ln(\xi + \mu). \tag{2.12}$$

An example of $w(\cdot)$ that corresponds to the shifted Gamma-distribution is depicted in Figure 3.

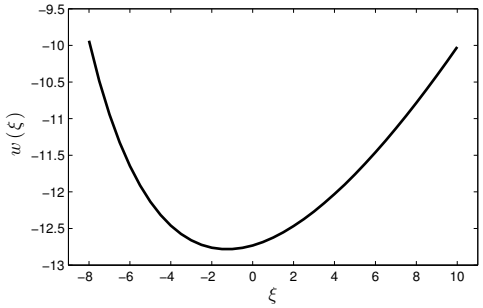

FIG. 3. *The function* $w(\xi)$ *defined in* (2.12)*, when* $\alpha = 10$*,* $\eta = 0.8$*, and* $\mu = \alpha/\eta$*.*

We then define $\boldsymbol{b}(t) := \boldsymbol{\psi}\big(\boldsymbol{a}(t)\big)$ as the *driving force* of energy transport or heat conduction,

$$\boldsymbol{\psi}(\boldsymbol{a}) := -\frac{\delta W(\boldsymbol{a})}{\delta \boldsymbol{a}}. \tag{2.13}$$

One can interpret $W(\boldsymbol{a})$ as the local entropy and $\boldsymbol{b}$ as the inverse temperature. It also relates the dynamics of $\boldsymbol{a}$ with its intrinsic distribution, which will be elaborated on in subsection 2.4.

We now propose extending the original MZ procedure by defining the projection

$$\mathcal{P}\boldsymbol{f} := \langle \boldsymbol{f}, \boldsymbol{b}^{\mathsf{T}} \rangle \langle \quad \rangle^{-1} \boldsymbol{b}, \tag{2.14}$$

where $\boldsymbol{b} := \boldsymbol{b}(0)$ for the sake of brevity and $\langle \cdot, \cdot \rangle$ stands for the covariance matrix at the canonical ensemble. When the energy is quadratic, or equivalently when $\boldsymbol{a}(t)$ follows a Gaussian distribution, the function $\boldsymbol{\psi}(\cdot)$ is linear, and the projection will be reduced to Mori's projection. Nevertheless, in general, these two projections are different. We will refer to this projection as an *entropy-based* projection.

Combined with (2.7), the formalism yields a generalized Langevin equation (GLE) written as

$$\dot{\boldsymbol{a}}(t) = \Omega \boldsymbol{b}(t) + \int_0^t \theta(t-s)\boldsymbol{b}(s)ds + \boldsymbol{F}(t), \tag{2.15}$$

where Ω is a constant matrix $\Omega = \langle \mathcal{L}\boldsymbol{a}(0), \boldsymbol{b}^{\mathsf{T}} \rangle \langle \boldsymbol{b}, \boldsymbol{b}^{\mathsf{T}} \rangle^{-1}$ and θ is the memory kernel $\theta(t) = \langle \mathcal{L}\boldsymbol{F}(t), \boldsymbol{b}^{\mathsf{T}} \rangle \langle \boldsymbol{b}, \boldsymbol{b}^{\mathsf{T}} \rangle^{-1}$. The first term in the general MZ equation (2.7) can be shown to be equal to zero due to symmetry if $\boldsymbol{a}$ is chosen as the local energy [13]. The GLE yields a memory term that is in a linear convolution form, while maintaining the nonlinearity of the model. As we will see in the next section, we use the GLE to directly describe the dynamics of the local energy, and it is quite straightforward to impose the random noise to ensure the correct statistics of the solution.

Incorporating fluctuations in the constitutive relation has already been suggested by Landau and Lifshitz [32]. Our viewpoint is along the same line as in the extended thermodynamics formalism (cf. Chapter 5) [29]. Namely, fluctuations in the constitutive relation, when supplemented to the energy equation, should lead to a stochastic model, rather than a deterministic one, which results in fluctuations in physical observables. Of our particular interest is the statistics of the energy (2.10), whose PDF is related to the entropy [22]: $W(\boldsymbol{a}) = \ln \rho(\boldsymbol{a}) + \text{const}$. The entropy extracted from the statistics can then be used to determine the deterministic part of the constitutive relation. Our projection formalism, together with the GLE (2.15), helps to reveal such subtle relations.

2.3. Local stochastic models via Markovian embedding. The GLE is nonlocal, which looks quite different from the conventional and extended models. To draw connections, we approximate the kernel function via its Laplace transform:

$$\Theta(\lambda) := \int_0^{+\infty} \theta(t) e^{-t/\lambda} dt. \tag{2.16}$$

We denote the statistics of the energy by

$$\begin{aligned} M(t) =& \langle \boldsymbol{a}(t), \boldsymbol{a}(0)^{\mathsf{T}} \rangle, \\ N(t) =& \langle \boldsymbol{b}(t), \boldsymbol{a}(0)^{\mathsf{T}} \rangle, \end{aligned} \tag{2.17}$$

whose Laplace transforms are denoted as $\widetilde{M}(\lambda)$ and $\widetilde{N}(\lambda)$,

$$\widetilde{M}(\lambda) = \int_0^{+\infty} M(t) e^{-t/\lambda} dt, \quad \widetilde{N}(\lambda) = \int_0^{+\infty} N(t) e^{-t/\lambda} dt. \tag{2.18}$$

Meanwhile, we assume that the short-time statistics can be extracted, e.g., from equilibrium MD simulations,

$$M_j = \langle \boldsymbol{a}^{(j)}(0), \boldsymbol{a}(0)^{\mathsf{T}} \rangle, \quad N_j = \langle \boldsymbol{b}^{(j)}(0), \boldsymbol{a}(0)^{\mathsf{T}} \rangle. \tag{2.19}$$

The superscripts j indicate the time derivative, $j \geq 0$. For example, we have that M_0 is the covariance matrix of $\boldsymbol{a}$. With direct calculations, one can also verify that $M_1 = 0,\ N_0 = -I$.

With these statistics, the Laplace transforms $\widetilde{M}(\lambda)$ and $\widetilde{N}(\lambda)$ can be expanded around 0_+ [5],

$$\begin{aligned} \widetilde{M}(\lambda) &= \lambda M_0 + \lambda^2 M_1 + \lambda^3 M_2 + \cdots, \\ \widetilde{N}(\lambda) &= \lambda N_0 + \lambda^2 N_1 + \lambda^3 N_2 + \cdots. \end{aligned} \tag{2.20}$$

For long-time statistics, we define

$$N_\infty = \lim_{\lambda \to +\infty} \widetilde{N}(\lambda) = \lim_{\varepsilon \to 0_+} \int_0^{+\infty} e^{-\varepsilon t} N(t) dt. \tag{2.21}$$

One systematic procedure for reducing the GLE is the embedding technique [35, 43]. As motivated by Mori's continued-fraction approach, we approximate $\Theta(\lambda)$ by rational functions,

$$R_{k,k} = \big[I - \lambda B_1 - \cdots - \lambda^k B_k\big]^{-1} \big[A_0 + \lambda A_1 + \cdots + \lambda^k A_k\big], \tag{2.22}$$

with coefficient matrices A_i's and B_i's to be determined by appropriate interpolation conditions.

Assuming that $\boldsymbol{a}(0)$ is uncorrelated with the noise term, we multiply the equation (2.15) by the transpose of $\boldsymbol{a}(0)$ and take the Laplace transform. With Θ approximated by $R_{k,k}$, we obtain

$$\frac{1}{\lambda}\widetilde{M}(\lambda) - M_0 = R_{k,k}(\lambda)\widetilde{N}(\lambda). \tag{2.23}$$

The coefficients in the rational function can be determined by matching the coefficients of λ's powers in the expansion around $\lambda = 0_+$ as follows:

$$\begin{aligned} M_1 &= A_0 N_0, \\ \lambda(M_2 - B_1 M_1) &= \lambda(A_0 N_1 + A_1 N_0), \\ \lambda^2(M_3 - B_1 M_2 - B_2 M_1) &= \lambda^2(A_0 N_2 + A_1 N_1 + A_2 N_0), \\ &\cdots. \end{aligned} \tag{2.24}$$

In general, one can match the first $2k$ coefficients, yielding $2k$ linear equations for the coefficients A_i's and B_i's. If $\boldsymbol{a}$ is the local energy, it is not difficult to verify that $N_{2k+1} = M_{2k+1} = 0$, which could further simplify the linear equations for A_i's and B_i's, e.g., $A_0 = 0$ [13]. The remaining condition is imposed at $\lambda \to +\infty$, which incorporates long-time statistics, yielding

$$M_0 = B_k^{-1} A_k N_\infty. \tag{2.25}$$

Overall, the interpolation is of Hermite type, with two interpolation points at $\lambda = 0_+$ and $\lambda = +\infty$. As it turns out, without the long-time statistics, the resulting model will be wave-type equations, with no dissipation.

Thanks to the rational approximation, the resulting reduced model can be converted back to the time domain and expressed as a set of differential equations. The memory term is embedded in an extended system which is Markovian (local).

2.4. A hierarchy of generalized heat conduction models. We present the models for $k = 0$, 1, and 2, to be referred to as the zeroth-, first-, and second-order models, respectively. An important problem in the implementation of the GLE is the approximation of the noise. Zhu and Venturi proposed using the Karhunen–Loève expansion and express the noise using the eigenmodes [71], while Chu and Li introduced Gaussian multiplicative noise to ensure the correct equilibrium statistics of CG variables [13]. In our approach the random noise in the GLE (2.15) will be approximated indirectly by introducing Gaussian additive noise in the extended system in such a way that the statistics of the local energy is consistent. We will also show that with additive Gaussian noise, matching two-point statistics is also straightforward. When substituting back into (2.15), this procedure leads to an approximation of $\boldsymbol{F}(t)$ using correlated Gaussian noise. The Gaussian nature of the noise from the MZ procedure can be verified with certain linear approximations of the atomic interactions [41] but otherwise remains an open issue. For example, approximations based on the more general Lèvy processes would be an interesting direction. With this approximation, the resulting stochastic equations can formally be interpreted as discretizations of stochastic partial differential equations (SPDEs).

Zeroth-order model. For $k = 0$, $\Theta(\lambda)$ is approximated by a constant matrix $R_{0,0} = \Gamma$. In order to accommodate long-term statistics, we pick an interpolating condition in (2.23) as λ goes to infinity, which yields

$$-M_0 = \Gamma N_\infty, \tag{2.26}$$

and

$$\Theta \approx \Gamma = -M_0 N_\infty^{-1}. \tag{2.27}$$

This corresponds to a delta function in the time domain, which is known as a Markovian approximation [26]. Since M_0 and N_∞ are associated with the statistics of energies in (2.19) and (2.21), Γ can be computed with the statistics of $\boldsymbol{a}$. Our numerical results suggest that Γ is a tridiagonal matrix and proportional to a three-point discrete Laplacian operator,

$$\Theta \approx -\kappa \nabla_h^2. \tag{2.28}$$

The subscript h is the width of each block. Here we use standard notation in finite-difference methods,

$$\nabla_h a_{j+1/2} := \frac{a_{j+1} - a_j}{h}, \quad \nabla_h \cdot q_j := \frac{q_{j+1/2} - q_{j-1/2}}{h}. \tag{2.29}$$

The subscript $j+1/2$ in the first equation indicates that the finite difference is operated on a_j and a_{j+1}, which are quantities defined at the center of the blocks. The result of the gradient operator is a quantity defined at the interfaces of adjacent blocks, here labeled by the index $j + 1/2$. Similarly, in the second equation, the divergence is operated on quantities at the interfaces, here denoted by $q_{j+1/2}$, e.g., the heat flux, and the result is a quantity defined at the block centers. The Laplacian operator in (2.28) is defined based on these two operators,

$$\nabla_h^2 a_j := \nabla_h \cdot \nabla_h a_j = \frac{a_{j+1} - 2a_j + a_{j-1}}{h^2}. \tag{2.30}$$

Therefore, the GLE (2.15) is reduced to

$$\dot{\boldsymbol{a}} = \kappa \nabla_h^2 \frac{\delta W(\boldsymbol{a})}{\delta \boldsymbol{a}} + \sigma \boldsymbol{\xi}. \tag{2.31}$$

The noise in the GLE (2.15) is approximated by standard Gaussian white noise $\boldsymbol{\xi}(t)$. The operator σ is not a scalar. Rather, it acts as a discrete divergence, which comes from the standard selection: if $\sigma\sigma^{\intercal} = -2\kappa\nabla_h^2$, then the equilibrium PDF (2.10) is the stationary PDF of our reduced model (2.31). Thus, $\nabla_h \cdot \xi(t)$ altogether can be interpreted as a discretized space-time white noise, and (2.31) can be viewed as a discrete analogue of a nonlinear heat equation driven by space-time white noise. The continuum limit for stochastic heat equations has been analyzed in [24, 20, 55]. Another remarkable observation is that this model shows similarity to the stochastic phase-field crystal model [4] as well as the general diffusion models [17, 19].

The local energy that we defined at the beginning satisfies the fundamental conservation law,

$$\dot{a}_j + \nabla_h \cdot q_j = 0, \tag{2.32}$$

with $q_{j+1/2}$ being the heat flux between the adjacent j and $j+1$ blocks. Explicit formulas of $\boldsymbol{q}$ can be derived for many MD models [10, 66].

Equations (2.31) and (2.32) suggest a stochastic constitutive relation for the heat flux,

$$q_{j+1/2} = -\kappa \nabla_h b_{j+1/2} + \sqrt{2\kappa}\xi_{j+1/2}, \quad b_j := \frac{\delta W(\boldsymbol{a})}{\delta a_j}, \tag{2.33}$$

where $\xi_{j+1/2}(t)$ is scalar standard Gaussian white noise. This implies a nonlinear, stochastic generalization of Fourier's law. It is also important in the constitutive modeling of heat flux to ensure the second law of thermodynamics, in that the corresponding entropy production rate remains nonnegative [29, 33]. It can be directly verified that the deterministic part of the nonlinear constitutive relation obeys this property, since $\nabla_h b_{j+1/2} \cdot \nabla_h a_{j+1/2} \geq 0$.

First-order model. When $k = 1$, the rational interpolating function becomes

$$R_{1,1}(\lambda) = \big[I - \lambda B_1\big]^{-1} \lambda A_1. \tag{2.34}$$

With this form of the approximation, the memory term can be greatly simplified. From the definition in (2.19), one can compute that $N_0 = -I$ and M_2 is proportional to a discrete Laplacian operator, i.e., $M_2 = \gamma_1 \nabla_h^2$. The matching conditions (2.24) lead to an explicit formula for the coefficient matrix A_1,

$$A_1 = -M_2 = -\gamma_1 \nabla_h^2. \tag{2.35}$$

Now, by incorporating the interpolation condition (2.25), we find that

$$B_1 = -\gamma_1/\kappa I. \tag{2.36}$$

To this end we define an auxiliary variable $\boldsymbol{z}$ and rewrite the GLE (2.15) as

$$\begin{cases} \dot{\boldsymbol{a}} = \boldsymbol{z}, \\ \dot{\boldsymbol{z}} = \gamma_1 \nabla_h^2 \dfrac{\delta W(\boldsymbol{a})}{\delta \boldsymbol{a}} - \dfrac{\gamma_1}{\kappa}\boldsymbol{z} + \sigma \boldsymbol{\xi}, \end{cases} \tag{2.37}$$

after plugging the parameters in. With the added noise, the variable $\boldsymbol{z}$ now embodies a *combined* approximation of both the memory and the random noise terms. To see this, one can apply the variation-of-constant formula and get

$$\begin{aligned}\boldsymbol{z}(t) = &\int_0^t \exp\left[-\tau\frac{\gamma_1}{\kappa}\right]\gamma_1\nabla_h^2\frac{\delta W(\boldsymbol{a})}{\delta \boldsymbol{a}}(t-\tau)d\tau \\ &+ \exp\left[-t\frac{\gamma_1}{\kappa}\right]\boldsymbol{z}(0) + \int_0^t \exp\left[-(t-\tau)\frac{\gamma_1}{\kappa}\right]\sigma d\boldsymbol{W}_\tau.\end{aligned} \tag{2.38}$$

The first term clearly corresponds to an approximation of the memory term, while the two remaining terms serve as an approximation of the random noise. The stationarity of this approximate noise can be verified using Itô isometry. See [43] for examples of such calculations.

By solving the Lyapunov equation [54], σ is given by the formula

$$\sigma\sigma^{\intercal} = -\frac{2\gamma_1^2}{\kappa}\nabla_h^2. \tag{2.39}$$

It is possible to select σ in other forms to ensure the stability, but by choosing σ in this way, the first-order model (2.37) has the stationary PDF,

$$\rho(\boldsymbol{a},\boldsymbol{z}) \propto \exp -\left[W(\boldsymbol{a}) - \frac{1}{2\gamma_1}\boldsymbol{z}^{\intercal}\nabla_h^{-2}\boldsymbol{z}\right], \tag{2.40}$$

which can be verified by the stationary Fokker–Planck equation. Clearly the marginal density associated with $\boldsymbol{a}$ is identical to the proposed PDF in (2.4). Again, together with the conservation law (2.32), this stochastic model (2.40) implies a constitutive relation for the heat flux,

$$\tau_1\dot{q}_{j+1/2} + q_{j+1/2} = -\kappa\nabla_h b_{j+1/2} + \sqrt{2\kappa}\xi_{j+1/2}, \quad b_j := \frac{\delta W(\boldsymbol{a})}{\delta a_j}. \tag{2.41}$$

Extended Thermodynamics

Rational Thermodynamics

This is an interesting generalization of the CV model [6]. Not only do we identify the origin of the relaxation parameter, $\tau_1 = \kappa/\gamma_1$, but we also incorporate a nonlinear driving force and a random noise. The model can also be written as

$$\tau_1\ddot{a}_j + \dot{a}_j = \kappa\nabla_h^2\left[\frac{\delta W(\boldsymbol{a})}{\delta \boldsymbol{a}}\right]_j + \tau_1\sigma\xi_j, \tag{2.42}$$

where σ behaves like a divergence operator $\nabla_h\cdot$ in light of (2.39). This equation corresponds to a discretization of a damped nonlinear wave equation with additive space-time white noise. However, unlike the zeroth-order (2.31), such a continuous limit is still an open issue even when the equations are linear. To the best of our knowledge, the existing finite element analysis only considered the case with white noise in time [16, 24, 67]. Therefore, the connection to an SPDE is only formal: By multiplying the SPDE by piecewise linear shape functions using the interpretation by Walsh [65], we move the divergence operator to the shape functions and obtain the semidiscrete form (2.42).

Second-order model. The procedure used in the first-order model can be extended to a higher order. When $k = 2$, we approximate $\theta(t)$ with $R_{2,2}(\lambda)$ in Laplace space based on (2.23). The second-order model can be expressed in a more compact form with an added noise term,

$$\dot{\boldsymbol{z}} = A\boldsymbol{b} + B\boldsymbol{z} + \sigma\boldsymbol{\xi}, \tag{2.43}$$

where A and B are block matrices,

$$A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \text{ and } B = \begin{bmatrix} B_1 & I \\ B_2 & 0 \end{bmatrix}. \tag{2.44}$$

To determine the covariance σ, we propose an equilibrium density of $(\boldsymbol{a}, \boldsymbol{z})$ in a separable form,

$$\rho(\boldsymbol{a}, \boldsymbol{z}) \propto \exp - \left[W(\boldsymbol{a}) + \frac{1}{2} \boldsymbol{z}^{\mathsf{T}} Q^{-1} \boldsymbol{z} \right]. \tag{2.45}$$

The marginal PDF of $\boldsymbol{a}$ is consistent with (2.10) by choosing Q according to the Lyapunov equation

$$\sigma\sigma^{\mathsf{T}} = -(BQ + QB^{\mathsf{T}}). \tag{2.46}$$

Such a procedure is standard [54], and it is based on solving the underlying Fokker–Planck equation.

When imposing the interpolation conditions to determine the parameters, this second-order approximation involves higher order statistical moments, M_4 and N_2. From numerical experiments, we observed that

$$M_4 \approx -\gamma_2 \nabla_h^2, \quad N_2 \approx -\kappa_1 \nabla_h^2, \tag{2.47}$$

but $\gamma_2 \gg \gamma_1 \kappa_1$. After direct substitutions, we obtain an explicit form for the reduced model (2.43),

$$\begin{cases} \dot{\boldsymbol{a}} = \boldsymbol{z}_1, \\ \dot{\boldsymbol{z}}_1 = \gamma_1 \nabla_h^2 \dfrac{\delta W(\boldsymbol{a})}{\delta \boldsymbol{a}} - \dfrac{\gamma_2 \kappa}{\gamma_1^2} \boldsymbol{z}_1 + \boldsymbol{z}_2, \\ \dot{\boldsymbol{z}}_2 = \dfrac{\gamma_2 \kappa}{\gamma_1} \nabla_h^2 \dfrac{\delta W(\boldsymbol{a})}{\delta \boldsymbol{a}} - \dfrac{\gamma_2}{\gamma_1} \boldsymbol{z}_1 + \sigma \boldsymbol{\xi}, \end{cases} \tag{2.48}$$

where $\sigma\sigma^{\mathsf{T}} = -2\frac{\gamma_2^2 \kappa}{\gamma_1^2} \nabla_h^2$. If we introduce a second relaxation parameter $\tau_2 = \gamma_1/\gamma_2$, the corresponding constitutive relation for the heat flux $\boldsymbol{q}$ reads as

$$\tau_2 \ddot{q}_{j+1/2} + \tau_1 \dot{q}_{j+1/2} + q_{j+1/2} = \kappa \nabla_h b_{j+1/2} + \frac{\tau_2 \kappa}{\tau_1} \nabla_h \dot{b}_{j+1/2} + \sqrt{2\kappa} \xi_{j+1/2}. \tag{2.49}$$

Here $b_j = -\frac{\delta W(\boldsymbol{a})}{\delta a_j}$ is defined according to (2.13). This is again a generalization of the heat conduction model known as the Tzou model [62], with parameters linked to the statistics of the local energies.

The procedure to embed the GLE (2.15) can be easily extended to higher orders, and it has been a common practice in reduced-order modeling. The details can be found in numerous publications [2, 35, 43]. For example, we have also derived a third-order model, and it can be regarded as a higher order relaxation model for the heat flux, but we omit the details here.

3. Numerical results. Our derivations yield stochastic constitutive relations up to any order. As with many data driven methods, the current approach works with the observation of the CG variables. The observations either can be obtained from experimental measurements or can be generated from direct numerical simulations.

However, often measured in nano-scale heat conduction is the local temperature, rather than the local energy. Therefore, in our numerical tests, we choose to run full molecular dynamics simulations to extract the time series of the local energy from the full trajectories. With partial observations of the local energies, we compute the projection direction $\boldsymbol{b}(\boldsymbol{a})$, which is related to the equilibrium distribution of CG variables $\boldsymbol{a}$, and determine the reduced models up to any order. The coefficients in these reduced models are identified by simple matrix algebra calculations of statistics of the CG variables without revisiting the underlying model.

Here, as validation, we solve the stochastic models (2.31), (2.37), (2.48), and a third-order model, and compare the numerical solutions with the true statistics obtained from full MD simulations.

For the reduced models, we will write all the reduced stochastic models in the following unified form, as suggested in (2.43):

$$\begin{aligned} \dot{\boldsymbol{a}} =& E_1 \boldsymbol{z}, \\ \dot{\boldsymbol{z}} =& A\boldsymbol{b} + B\boldsymbol{z} + \sigma\boldsymbol{\xi}. \end{aligned} \tag{3.1}$$

Here, $E_1 = (I, 0, \ldots, 0)$ singles out $\boldsymbol{z}_1$ from the auxiliary variables $\boldsymbol{z}$. Then, the integrator is constructed by splitting the equations into two systems,

$$\begin{aligned} &\text{(I)} \qquad \dot{\boldsymbol{a}} = E_1\boldsymbol{z}, \qquad \dot{\boldsymbol{z}} = \boldsymbol{0}, \\ &\text{(II)} \qquad \dot{\boldsymbol{a}} = \boldsymbol{0}, \qquad \dot{\boldsymbol{z}} = A\boldsymbol{b} + B\boldsymbol{z} + \sigma\boldsymbol{\xi}, \end{aligned} \tag{3.2}$$

both of which are linear by themselves, and the solutions can be written out explicitly. A second-order weak method can be constructed by solving the first equation for half of the step, the second equation for one step, and then the first equation for half of the step.

More specifically, within one time from t_n to t_{n+1}, the algorithm consists of the following steps:

$$\begin{aligned} \boldsymbol{a}_{n+1/2} &\leftarrow \boldsymbol{a}_n + E_1\boldsymbol{z}_n\Delta t/2, \\ \boldsymbol{z}_{n+1} &\leftarrow C_0\boldsymbol{z}_n + C_1 A\boldsymbol{b}_{n+1/2} + \Delta W_n, \\ \boldsymbol{a}_{n+1} &\leftarrow \boldsymbol{a}_{n+1/2} + E_1\boldsymbol{z}_{n+1}\Delta t. \end{aligned} \tag{3.3}$$

The coefficients can be found by solving the second equation in (3.2) using a variation-of-constants formula similar to (2.38):

$$\begin{aligned} \boldsymbol{z}(t_{n+1}) =& \exp\big[\Delta t B\big]\boldsymbol{z}(t_n) + \int_0^{\Delta t} \exp\big[(\Delta t - \tau)B\big] A\boldsymbol{b}(\boldsymbol{a}) d\tau \\ &+ \int_0^{\Delta t} \exp\big[(\Delta t - \tau)B\big]\sigma dW_\tau. \end{aligned} \tag{3.4}$$

In the integral, $\boldsymbol{a}$ actually remains a constant. Therefore, the integral can be explicitly computed. The last term is Gaussian process with zero mean, and the variance can be computed using the Itô isometry. With direct calculations, we find that

$$C_0 = \exp\big[\Delta t B\big], \quad C_1 = B^{-1}\big(\exp\big[\Delta t B\big] - I\big), \tag{3.5}$$

and ΔW_n is a normal random variable with mean zero and variance Σ given by

$$\Sigma = Q - \exp\big[\Delta t B\big] Q \exp\big[\Delta t B^T\big]. \tag{3.6}$$

As far as the weak convergence is concerned, this can be understood in terms of the one-step transition density, which is unknown in general for the underlying Fokker–Planck equation. The operator-splitting scheme approximates the transition density using two stochastic differential equations (SDEs), for which the transition kernels are computable. Since this is by now a rather standard method in molecular simulations, we will refer the reader to the monographs for the explicit formulas [37, 36].

As a simple validation of the choices of the coefficients of the noise, we compared the PDF obtained from the models (2.31) and (2.37) to that from the full MD simulation. As shown in Figure 4, the reduced models reproduce the correct non-Gaussian statistics.

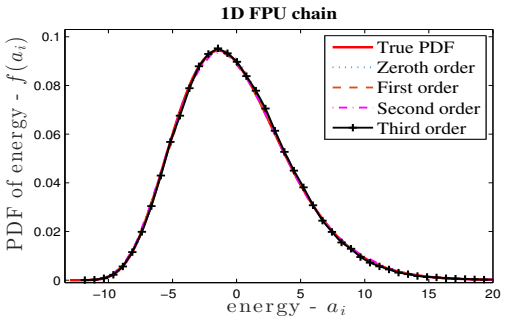

FIG. 4. *PDFs of a_5 in 1D FPU chain example. Histograms of full MD simulations (True PDF) and histograms of solutions to reduced models (Zeroth, First, Second, Third orders) are compared. The time step is 0.001a.u. for reduced model simulations, and the Milstein method is used as the stochastic integration scheme.*

We compute the two-point time correlations of the energy to examine dynamical properties,

$$C_i(t) = \langle a_i(t), a_i(0) \rangle. \tag{3.7}$$

As shown in Figure 5, the two-point statistics of the energy is consistent for both the 1D chain model (left) and the nanotube system (right). For both systems, the zeroth-order approximation (2.31), which is a nonlinear extension of Fourier's law, exhibits large error. We observed improved accuracy as we increase the order of the approximations. Interestingly, the 1D model seems to be a bit harder to approximate: We had to extend the method to the third order to obtain a good approximation of the time correlation. On the other hand, for the nanotube system, the second-order model (2.37) is able to offer quite satisfactory results within our simulation window. The explanation for this is related to the central limit theorem. With the same block size, there are more atoms in the nanotube system. The statistics of the local energy is less skewed and tends to be more Gaussian. So a low order model is already able to capture statistics.

4. Summary and discussions. In summary, we derived generalized heat conduction models from the underlying MD model. The stochastic constitutive equations can be nonlinear, and the parameters were linked directly to the statistics of the local energy, making it possible to determine system-specific model parameters using statistical properties. The nonlinearity can be attributed to the non-Gaussian statistics of the local energies. With approximations of the memory integral, we obtained various models for the heat flux, which can be viewed as nonlinear and stochastic extensions

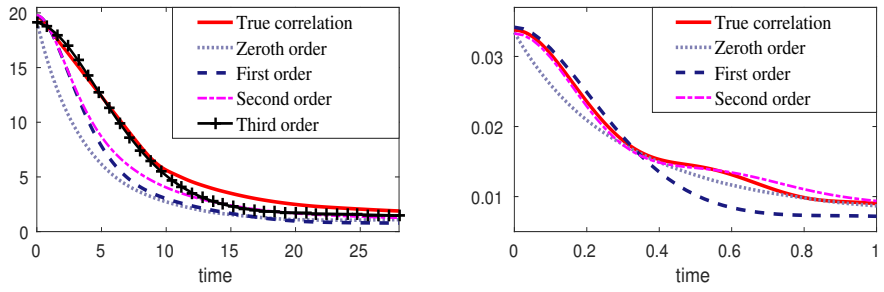

FIG. 5. *Two-point statistics of a_1 from the full MD simulations (True correlation) and reduced models. Left: 1D FPU chain model. Right: Single-wall nanotube. As the model order increases, the plots demonstrate convergence to the true time correlation.*

of the Cattaneo–Vernotte and Guyer–Krumhansl relaxation models. The models were validated by examining the one- and two-point statistics.

Our parameter identification procedure uses statistics, which is made possible by the stochastic nature of the models. This is quite different from the parameter identification methods based on deterministic models [56].

Our effort to develop stochastic models was motivated by the classical text of Landau, Lifshitz, and Pitaevskii [32], the recent stochastic parameter estimation method [35], and many experimental observations of thermal fluctuations in diffusion processes. This approach can be extended to higher dimensions, e.g., graphene sheet, as well as other diffusion problems, e.g., ion diffusions. These works are underway.

REFERENCES

[1] F. ALVAREZ AND D. JOU, *Memory and nonlocal effects in heat transport: From diffusive to ballistic regimes*, Appl. Phys. Lett., 90 (2007), 083109.

[2] Z. BAI, P. M. DEWILDE, AND R. W. FREUND, *Reduced-order modeling*, in Handbook of Numerical Analysis, Vol. XIII, Hanb. Numer. Anal. 13, North-Holland, Amsterdam, 2005, pp. 825–891.

[3] A. A. BALANDIN, *Thermal properties of graphene and nanostructured carbon materials*, Nat. Mater., 10 (2011), 569.

[4] J. BERRY, K. ELDER, AND M. GRANT, *Melting at dislocations and grain boundaries: A phase field crystal study*, Phys. Rev. B, 77 (2008), 224114.

[5] N. BLEISTEIN AND R. A. HANDELSMAN, *Asymptotic Expansions of Integrals*, Dover, New York, 1986.

[6] C. CATTANEO, *A form of heat-conduction equations which eliminates the paradox of instantaneous propagation*, Comptes Rendus, 247 (1958), pp. 431–433.

[7] C. W. CHANG, D. OKAWA, H. GARCIA, A. MAJUMDAR, AND A. ZETTL, *Breakdown of Fourier's Law in nanotube thermal conductors*, Phys. Rev. Lett., 101 (2008), 075903.

[8] J. CHEN, G. ZHANG, AND B. LI, *Molecular dynamics simulations of heat conduction in nanostructures: Effect of heat bath*, J. Phys. Soc. Japan, 79 (2010), 074604.

[9] M. CHEN, X. LI, AND C. LIU, *Computation of the memory functions in the generalized Langevin models for collective dynamics of macromolecules*, J. Chem. Phys., 141 (2014), 064112.

[10] Y. CHEN, *Local stress and heat flux in atomistic systems involving three-body forces*, J. Chem. Phys., 124 (2006), 054113.

[11] A. J. CHORIN, O. H. HALD, AND R. KUPFERMAN, *Optimal prediction and the Mori-Zwanzig representation of irreversible processes*, Proc. Natl. Acad. Sci. USA, 97 (2000), pp. 2968–2973.

[12] A. J. CHORIN AND P. STINIS, *Problem reduction, renormalization, and memory*, Commun. Appl. Math. Comput. Sci., 1 (2006), pp. 1–27.

[13] W. CHU AND X. LI, *The Mori-Zwanzig formalism for the derivation of a fluctuating heat conduction model from molecular dynamics*, Commun. Math. Sci., 17 (2019), pp. 539–563.

[14] G. Ciccotti and J.-P. Ryckaert, *On the derivation of the generalized Langevin equation for interacting Brownian particles*, J. Stat. Phys., 26 (1981), pp. 73–82.
[15] E. Darve, J. Solomon, and A. Kia, *Computing generalized Langevin equations and generalized Fokker-Planck equations*, Proc. Natl. Acad. Sci. USA, 106 (2009), pp. 10884–10889.
[16] A. Davie and J. Gaines, *Convergence of numerical schemes for the solution of parabolic stochastic partial differential equations*, Math. Comp., 70 (2001), pp. 121–134.
[17] J. De La Torre, P. Español, and A. Donev, *Finite element discretization of non-linear diffusion equations with thermal fluctuations*, J. Chem. Phys., 142 (2015), 094115.
[18] A. Dhar, *Heat transport in low-dimensional systems*, Adv. Phys., 57 (2008), pp. 457–537.
[19] A. Donev, E. Vanden-Eijnden, A. Garcia, and J. Bell, *On the accuracy of finite-volume schemes for fluctuating hydrodynamics*, Commun. Appl. Math. Comput. Sci., 5 (2010), pp. 149–197.
[20] Q. Du and T. Zhang, *Numerical approximation of some linear stochastic partial differential equations driven by special additive noises*, SIAM J. Numer. Anal., 40 (2002), pp. 1421–1445, https://doi.org/10.1137/S0036142901387956.
[21] P. Español, *Statistical mechanics of coarse-graining*, in Novel Methods in Soft Matter Simulations, Springer, New York, 2004, pp. 69–115.
[22] G. Faure, R. Delgado-Buscalioni, and P. Español, *The entropy of a complex molecule*, J. Chem. Phys., 146 (2017), 224106.
[23] R. A. Guyer and J. Krumhansl, *Solution of the linearized phonon Boltzmann equation*, Phys. Rev., 148 (1966), 766.
[24] I. Gyöngy, *Lattice approximations for stochastic quasi-linear parabolic partial differential equations driven by space-time white noise* II, Potential Analysis, 11 (1999), pp. 1–37.
[25] C. Hijón, P. Español, E. Vanden-Eijnden, and R. Delgado-Buscalioni, *Mori–Zwanzig formalism as a practical computational tool*, Faraday Discuss., 144 (2010), pp. 301–322.
[26] C. Hijón, M. Serrano, and P. Español, *Markovian approximation in a coarse-grained description of atomic systems*, J. Chem. Phys., 125 (2006), 204101.
[27] J. Hu, X. Ruan, and Y. P. Chen, *Thermal conductivity and thermal rectification in graphene nanoribbons: A molecular dynamics study*, Nano Lett., 9 (2009), pp. 2730–2735.
[28] S. Izvekov and G. A. Voth, *Modeling real dynamics in the coarse-grained representation of condensed phase systems*, J. Chem. Phys., 125 (2006), pp. 151101–151104.
[29] D. Jou, J. Casas-Vázquez, and G. Lebon, *Extended irreversible thermodynamics*, in Extended Irreversible Thermodynamics, Springer, New York, 1996, pp. 41–74.
[30] D. Kauzlarić, J. T. Meier, P. Español, A. Greiner, and S. Succi, *Markovian equations of motion for non-Markovian coarse-graining and properties for graphene blobs*, New J. Phys., 15 (2013), 125015.
[31] A. J. Ladd, B. Moran, and W. G. Hoover, *Lattice thermal conductivity: A comparison of molecular dynamics and anharmonic lattice dynamics*, Phys. Rev. B, 34 (1986), 5058.
[32] L. D. Landau, E. M. Lifshitz, and L. Pitaevskii, *Statistical Physics, Part* I, Elsevier, Amsterdam, 1980.
[33] G. Lebon, *Heat conduction at micro and nanoscales: A review through the prism of extended irreversible thermodynamics*, J. Non-Equilibrium Thermodynamics, 39 (2014), pp. 35–59.
[34] J.-U. Lee, D. Yoon, H. Kim, S. W. Lee, and H. Cheong, *Thermal conductivity of suspended pristine graphene measured by Raman spectroscopy*, Phys. Rev. B, 83 (2011), 081419.
[35] H. Lei, N. A. Baker, and X. Li, *Data-driven parameterization of the generalized Langevin equation*, Proc. Natl. Acad. Sci. USA, 113 (2016), pp. 14183–14188.
[36] B. Leimkuhler and C. Matthews, *Molecular Dynamics*, Springer, New York, 2002.
[37] B. Leimkuhler and C. Matthews, *Molecular Dynamics*, Interdiscip. Appl. Math. 39, Springer, Cham, 2015.
[38] S. Lepri, R. Livi, and A. Politi, *Heat conduction in chains of nonlinear oscillators*, Phys. Rev. Lett., 78 (1997), 1896.
[39] S. Lepri, R. Livi, and A. Politi, *On the anomalous thermal conductivity of one-dimensional lattices*, Europhys. Lett., 43 (1998), 271.
[40] S. Lepri, R. Livi, and A. Politi, *Thermal conduction in classical low-dimensional lattices*, Phys. Rep., 377 (2003), pp. 1–80.
[41] X. Li, *A coarse-grained molecular dynamics model for crystalline solids*, Internat. J. Numer. Methods Engrg., 83 (2010), pp. 986–997.
[42] Z. Li, X. Bian, X. Li, and G. E. Karniadakis, *Incorporation of memory effects in coarse-grained modeling via the Mori-Zwanzig formalism*, J. Chem. Phys., 143 (2015), 243128.
[43] L. Ma, X. Li, and C. Liu, *The derivation and approximation of coarse-grained dynamics from Langevin dynamics*, J. Chem. Phys., 145 (2016), 204117.
[44] S. Maruyama, *A molecular dynamics simulation of heat conduction in finite length SWNTs*,

Phys. B, 323 (2002), pp. 193–195.
[45] A. J. H. McGaughey and M. Kaviany, *Thermal conductivity decomposition and analysis using molecular dynamics simulations. Part* I. *Lennard-Jones argon*, Int. J. Heat Mass Transfer, 47 (2003), pp. 1783–1798.
[46] N. Mingo and D. Broido, *Carbon nanotube ballistic thermal conductance and its limits*, Phys. Rev. Lett., 95 (2005), 096105.
[47] H. Mori, *A continued-fraction representation of the time-correlation functions*, Progr. Theoret. Phys., 34 (1965), pp. 399–416.
[48] H. Mori, *Transport, collective motion, and Brownian motion*, Progr. Theoret. Phys., 33 (1965), pp. 423–450.
[49] S. Nosé, *A molecular dynamics method for simulations in the canonical ensemble*, Mol. Phys., 52 (1984), pp. 255–268.
[50] E. J. Parish and K. Duraisamy, *Non-Markovian closure models for large eddy simulations using the Mori-Zwanzig formalism*, Phys. Rev. Fluids, 2 (2017), 014604.
[51] H. Pascal, *A nonlinear model of heat conduction*, J. Phys. A Math. Gen., 25 (1992), 939.
[52] C. Prévôt and M. Röckner, *A Concise Course on Stochastic Partial Differential Equations*, Lecture Notes in Math. 1905, Springer, Berlin, 2007.
[53] J. Price and P. Stinis, *Renormalized reduced order models with memory for long time prediction*, Multiscale. Model. Simul., 17 (2019), pp. 68–91, https://doi.org/10.1137/17M1151389.
[54] H. Risken and H. Haken, *The Fokker-Planck Equation: Methods of Solution and Applications*, 2nd ed., Springer, New York, 1989.
[55] K.-U. Schaumlöffel, *White noise in space and time and the cylindrical Wiener process*, Stochastic Anal. Appl., 6 (1988), pp. 81–89.
[56] A. Singh and E. B. Tadmor, *Thermal parameter identification for non-Fourier heat transfer from molecular dynamics*, J. Comput. Phys., 299 (2015), pp. 667–686.
[57] H. Spohn, *Nonlinear fluctuating hydrodynamics for anharmonic chains*, J. Stat. Phys., 154 (2014), pp. 1191–1227.
[58] M. Stepanova, *Dynamics of essential collective motions in proteins: Theory*, Phys. Rev. E, 76 (2007), 051918.
[59] J. Tersoff, *New empirical model for the structural properties of silicon*, Phys. Rev. Lett., 56 (1986), pp. 632–635.
[60] M. Toda, R. Kubo, and N. Hashitsume, *Statistical Physics* II. *Nonquilibrium Statistical Mechanics*, Springer, New York, 1983.
[61] M. E. Tuckerman, B. J. Berne, and G. J. Martyna, *Reversible multiple time scale molecular dynamics*, J. Chem. Phys., 97 (1990), pp. 1990–2001.
[62] D. Tzou, *Nonlocal behavior in phonon transport*, Int. J. Heat Mass Transfer, 15 (2011), p. 475–481.
[63] P. Vernotte, *Paradoxes in the continuous theory of the heat equation*, C.R. Acad. Sci., 246 (1958), pp. 154–3.
[64] S. G. Volz and G. Chen, *Molecular dynamics simulation of thermal conductivity of silicon nanowires*, Appl. Phys. Lett., 75 (1999), pp. 2056–2058.
[65] J. B. Walsh, *An Introduction to Stochastic Partial Differential Equations*, Springer, New York, 1986.
[66] X. Wu and X. Li, *On consistent definitions of momentum and energy fluxes for molecular dynamics models with multi-body interatomic potentials*, Model. Simul. Mater. Sci. Eng., 23 (2015), 015003.
[67] Y. Yan, *Semidiscrete Galerkin approximation for a linear stochastic parabolic partial differential equation driven by an additive noise*, BIT, 44 (2004), pp. 829–847.
[68] H. Yoshida, *Construction of higher order symplectic integrators*, Phys. Lett. A, 150 (1990), pp. 262–268.
[69] Y. Zhu, J. M. Dominy, and D. Venturi, *On the estimation of the Mori-Zwanzig memory integral*, J. Math. Phys., 59 (2018), 103501.
[70] Y. Zhu and D. Venturi, *Faber approximation of the Mori-Zwanzig equation*, J. Comput. Phys., 372 (2018), pp. 694–718.
[71] Y. Zhu and D. Venturi, *Generalized Langevin Equations for Systems with Local Interactions*, preprint, https://arxiv.org/abs/1906.04918, 2019.
[72] R. Zwanzig, *Nonlinear generalized Langevin equations*, J. Stat. Phys., 9 (1973), pp. 215–220.