COMMUN. MATH. SCI.
Vol. 17, No. 2, pp. 539–563

THE MORI-ZWANZIG FORMALISM FOR THE DERIVATION OF A FLUCTUATING HEAT CONDUCTION MODEL FROM MOLECULAR DYNAMICS*

WEIQI CHU† AND XIANTAO LI‡

Abstract. Energy transport equations are derived directly from a many-particle system as a coarse-grained (CG) description. This effort is motivated by the observation that the conventional heat equation is unable to describe the heat conduction process at the nano-mechanical scale. With the local energy density chosen as the CG variables, we apply the Mori-Zwanzig formalism to derive a reduced model, in the form of a generalized Langevin equation. A Markovian embedding technique is then employed to eliminate the history dependence. Meanwhile, auxiliary variables are introduced to establish auxiliary equations that govern the dynamics of the energy flux. In sharp contrast to conventional energy transport models, this derivation yields stochastic dynamical models for the spatially averaged energy. The random force in the generalized Langevin equation is typically modeled by additive white Gaussian noise. As an initial attempt, we consider multiplicative white Gaussian noise, to ensure the correct statistics of the non-Gaussian solution.

Keywords. Molecular dynamics; Nanoscale Energy Transport; Mori-Zwanzig formalism.

AMS subject classifications. 60H10; 65C20.

1. Introduction

During the past two decades, there has been a rapidly growing interest in modeling heat transport at the microscopic scale [8, 26, 36, 57, 58]. Such renewed interest has been driven by the progress in designing and manufacturing micro mechanical and electrical devices, for which thermal conduction properties have great influences on the performance and reliability. As the size of electrical and mechanical devices is decreased to the micron and sub-micron scales, they often exhibit heat conduction properties that are quite different from the observations at the macroscopic level. For example, there has been overwhelming evidence indicating the failure of the conventional Fourier's law (and therefore the standard heat equation). Furthermore, heat pulses were observed in experiments [96], which are typical behavior of wave equations.

From a modeling viewpoint, a natural approach to incorporate (and predict) some of the observed behavior is to modify the traditional heat equation, *e.g.*, by introducing nonlocal terms or higher order derivatives [96, 97]. These generalized models have been successful in interpreting heat pulse propagation and they are easy to work with, but they are quite ad hoc. For example, such a generalization inevitably introduces additional model parameters that are difficult to determine.

Thus far, the most popular approach to study heat conduction is direct molecular dynamics (MD) simulations, which are often designed to mimic experimental setup. Given an interatomic potential V, either empirically constructed or derived from more fundamental considerations, an MD model is typically expressed in terms of Newton's equations of motion. There are many well established computational techniques for MD

*Received: March 27, 2018; Accepted (in revised form): December 27, 2018. Communicated by Arnulf Jentzen.
Research supported by the National Science Foundation DMS-1522617, DMS-1619661 and DMS-1819011.

†Department of Mathematics, Pennsylvania State University, University Park, PA 16802-6400, USA (wzc122@psu.edu).

‡Department of Mathematics, Pennsylvania State University, University Park, PA 16802-6400, USA (xli@math.psu.edu).

simulations [2,33]. They have motivated a lot of recent efforts to understand the origin and limitations of Fourier's Law [9, 10, 36, 57–59]. Furthermore, most of the studies have focused on the dependence of the heat conductivity on the geometry, length and temperature of the system. A typical setup is to connect the boundaries to two heat baths with different temperatures, modeled by stochastic (Langevin) or deterministic (Nosé-Hoover [75]) thermostats. The MD equations are solved to drive the system to a steady state, at which point the heat flux can be measured to estimate the heat conductivity. More general transient heat conduction problems, however, would require more substantial efforts.

Nevertheless, in spite of the many contributions that have recently appeared on the studies of heat conduction processes (*e.g.*, [14,15,26,37,42,47,68,69,78,88,99,101–103]), direct MD models have several serious limitations when applied to heat conduction problems. The first obvious limitation is the computational cost. One has to work with a large number of atoms, and the system needs to be integrated for millions or billions of steps. As a consequence, most current MD studies are restricted to quasi-one-dimensional systems, *e.g.*, nanowires [16, 17, 23, 26, 60, 65, 100, 104], nanotubes [13, 35, 44, 51, 67, 76, 78, 99, 107], and nanoribbons [31, 45, 87].

Further, while it is often straightforward to incorporate quantities such as displacement, velocity, temperature, and pressure into MD simulations as constraints, however, the temperature gradient is very difficult to impose. The temperature gradient that can be imposed is usually on the order of $10^8 - 10^9 K/m$, which is too large to model realistic systems. It is unclear whether results obtained in such MD simulations can be appropriately extrapolated to the correct regime.

This paper is strongly motivated by the above-mentioned issues, and the purpose is to present a coarse-grained (CG) model to alleviate these fundamental modeling difficulties. The CG procedure drastically reduces the number of degrees of freedom and offers a practical alternative. Coarse-graining methodologies have found enormous applications in material science problems and biological problems [5,21,22,38–40,46,49, 61,72–74,79–81,84–86,89,90,93,108]. Many CG models have been developed and they have shown great promise in reducing the computational cost and efficiently capturing the primary quantities of interest. However, most existing CG molecular models are focused on finding the effective potentials, known as the potential of mean forces, at a *constant* temperature. These existing CG models are in the similar form as the MD models with a possible addition of damping terms or random forces and typically take the CG variables as the linear combination of positions and velocities, describing only the time evolution of the averaged position and momentum.

Our approach aims at energy transport process, and we pick the CG variables to be the local energy, which is in general a nonlinear function. In order to derive from the many-particle dynamics a physical model in the ambient space, we start by defining a locally averaged internal energy as CG variables, and we use the Mori-Zwanzig (MZ) formalism [71, 109] to first derive an *exact* equation for these variables. In particular, we choose Mori's orthogonal projection [71] to project the equations to the subspace spanned by the CG variables. For such CG variables, this projection yields a memory term, which exhibits a simple form of a convolution in time. To alleviate the effort to compute the memory term at every step, we use the Markovian embedding techniques, recently developed in [12,56,66], to approximate the memory using an extended system of differential equations with *no* memory. The additional equations govern the dynamics of the energy flux, and they will only reduce to the conventional Fourier's law in a limiting case.

The idea of using such a generalized constitutive model dates back as early as in 1950's. For instance, Cattaneo and Vernotte have proposed a relaxation model for the heat flux [11,98], which unlike the traditional heat equation, admits thermal waves with finite propagation speed. Later, Guyer and Krumhansl [41] derived higher order models from the linearized phonon Boltzmann equation by using asymptotic expansions. The application of the MZ formalism to derive hydrodynamics equations, including energy transport, has been discussed in [1], and the extended system can be expressed in a continued-fraction form. Jou and co-workers formulated a unified framework and established the principle of extended irreversible thermodynamics [3, 48], where the entropy is generalized to include fluxes which led to additional thermodynamic relations. Determining the coefficients for such extended system is usually a nontrivial task. Singh and Tadmor [91] proposed a parameter identification algorithm using cosine and sine signals.

We propose to determine the parameters from the statistics of the underlying atomistic description, which would take into account the crystal structure, the size of the system, and the detailed interactions between the atoms. In particular, the parameters are obtained through a Hermite interpolation of the Laplace transform of the time correlation function, where the function values are directly linked to the statistics of the CG variables, which can be extracted from an equilibrium MD simulation (or possibly, from experimental measurements).

In principle, the noise term can be averaged out by simply taking the average of every term in the CG model. This is a particular advantage of the Mori's projection [21]. However, motivated by the crucial observation that many mechanical systems at the micron scale or smaller are subject to strong fluctuations, we will forgo such an averaging step, and work with the models *with* the random noise. This results in an energy transport model *with fluctuations*, represented by a system of stochastic differential equations (SDE). This is in the same spirit as the fluctuating hydrodynamics models [7, 27]. Consequently, the solutions are expected to be stochastic in nature. An important issue naturally arises: How does one guarantee that the corresponding solution has the correct statistics? This is an issue that has not been thoroughly investigated in the literature. The second half of this paper will be focused on the approximation of the noise. We first consider an approximation of the random force by an additive Gaussian white noise, in which case the solution should have Gaussian statistics. Unfortunately, by examining a one-dimensional chain model, we have found that the correct statistics behave more like a Gamma distribution. In particular, the energy must have a lower bound. Although the approximation by an additive noise yields reasonable approximations to the time correlations, the probability density function (PDF) of the solution is incorrect.

To ensure that the correct PDF is obtained, we make an attempt to approximate the noise by a *multiplicative* noise. In this case, the diffusion constant depends on the solution itself. We determine the diffusion coefficient by solving the steady-state Fokker-Planck equation. We are able to find a diagonal matrix for the diffusion coefficient matrix such that the Gamma distribution is an equilibrium probability density. As a further extension, we introduce a higher order approximation where both the CG energy variables and their time derivatives have the correct PDF. This leads to a Langevin type of equation with multiplicative noise. To the best of our knowledge, such models have not been reported in the literature.

We point out that existing stochastic heat conduction models have been proposed by Español et al. in [4,28,82], as extensions of the dissipative particle dynamics (DPD) [30].

Other relevant efforts include the recent work in deriving non-isothermal CG models of energy transport [29], which is an extension of the dissipative particle dynamics with energy conservation (DPDE) model.

The rest of the paper is organized as follows: In Section 2, we discuss the mathematical derivation and examine the general properties of the generalized Langevin equations derived from the Mori's projection. We introduce the Markovian embedding technique for the approximation of the memory term. Then in Section 3, we present the approximation of the random noise using a one-dimensional system as an example.

2. The derivation and numerical approximations

2.1. The general projection formalism. Our starting point is an all-atom description, which embodies the detailed interactions among all the atoms in the system. More specifically, let x and v be position and velocity of the atoms respectively; $x,v\in\mathbb{R}^{dN}$ with d being the space dimension and N being the total number of atoms. The dynamics follows Newton's second law,

$$\begin{cases} \dot{x}=v, & x(0)=x_0,\\ m\dot{v}=f(x)=-\dfrac{\partial V(x)}{\partial x}, & v(0)=v_0, \end{cases} \tag{2.1}$$

where $V(x)$ is the potential energy of the system.

It is important for our derivation to include the dependence on the initial configuration. Let $y_0=(x_0,v_0)$ be the initial state, we write solutions of (2.1) as $x(y_0,t)$ and $v(y_0,t)$ to explicitly indicate the dependence on y_0. The initial condition will be selected based on a probability measure μ_0, assumed to have a probability density function given by $\rho_0(y_0)$.

Next, we consider a quantity of interest (QOI) a, defined by an n-dimensional function $\psi(x,v)$,

$$a(y_0,t)=\psi(x(y_0,t),v(y_0,t)). \tag{2.2}$$

In statistical mechanics, $a(y_0,t)$ is often referred to as CG variable. For a short notation, we denote a without parenthesis as the initial value, *i.e.*, $a=a(y_0,0)$.

Further, we assume that the CG variables belong to a Hilbert space $L^2(\cdot,d\mu_0)$, the L^2 inner-product space weighted by the density ρ_0. We define the average $\langle\cdot\rangle$ and the correlation matrix $\langle\cdot,\cdot\rangle$ on $L^2(\cdot,d\mu_0)$ component-wisely, as follows

$$\begin{aligned} \langle a\rangle_i(t) &:= \int_\Psi a_i(y_0,t)\rho_0(y_0)dy_0, \quad 1\le i\le n,\\ \langle a,b\rangle_{ij}(t) &:= \int_\Psi a_i(y_0,t)b_j(y_0,t)\rho_0(y_0)dy_0, \quad 1\le i,j\le n, \end{aligned} \tag{2.3}$$

where $\Psi=\mathbb{R}^{dN}$ is the phase space. Here ρ_0 is the initial configuration density function of (2.1). The most popular choice is the canonical ensemble, for example, by Mori [70]. But in principal, ρ_0 can be more general, and even out-of-equilibrium. Several choices have been suggested by Zwanzig [110].

The goal is to derive a reduced equation for the CG variables $a(y_0,t)$. In this paper, we follow the MZ procedure [19–21,71,109], and define the propagating operator $\mathcal{L}$ as,

$$\mathcal{L}:=v_0\cdot\frac{\partial}{\partial x_0}+\frac{f(x_0)}{m}\cdot\frac{\partial}{\partial v_0}. \tag{2.4}$$

For the CG variables $a \in L^2(\cdot, d\mu_0)^n$, one has the formal relation

$$\begin{aligned} a(\cdot,t) &= e^{t\mathcal{L}} a(\cdot,0), \\ \dot{a}(\cdot,t) &:= \frac{\partial a(\cdot,t)}{\partial t} = \mathcal{L} e^{t\mathcal{L}} a(\cdot,0), \end{aligned} \tag{2.5}$$

where $e^{t\mathcal{L}}$ is the Koopman operator [52] that propagates an observable.

A key step in the MZ formulation is a projection operator $\mathcal{P}$ that maps functions to the space spanned by a. We adopt the orthogonal projection suggested by Mori [71].

For each function $b \in L^2(\cdot, d\mu_0)^n$, Mori's projection $\mathcal{P}$ is defined as follows,

$$\mathcal{P}b := \langle b, a \rangle M^{-1} a, \tag{2.6}$$

where M^{-1} is the inverse of matrix $M = \langle a, a \rangle$. We also define $\mathcal{Q}$ as the complementary operator of $\mathcal{P}$, *i.e.*, $\mathcal{Q} = \mathcal{I} - \mathcal{P}$.

REMARK 2.1. Note that the covariance matrix M only involves the one-point statistics of a and can be guaranteed to be nonsingular by carefully selecting the CG variables. In practice, this corresponds to the appropriate choice of ψ so that the CG variables are not redundant. Even in the case when the CG variables are redundant, *e.g.*, when energy is conservative and the matrix M becomes singular, the projection can still be well defined by interpreting M^{-1} as the pseudo-inverse.

Once the projection operator is in place, the Mori-Zwanzig formalism can be invoked, and the following generalized Langevin equation (GLE) can be derived [71],

$$\dot{a}(\cdot,t) = \Omega a(\cdot,t) - \int_0^t \theta(t-s) a(\cdot,s) ds + F(\cdot,t), \tag{2.7}$$

where

$$\Omega = \langle \mathcal{L}a, a \rangle M^{-1}, \; F(\cdot,t) = e^{t\mathcal{Q}\mathcal{L}} \mathcal{Q}\mathcal{L} a, \text{ and } \theta(t) = -\langle \mathcal{L}F(t), a \rangle M^{-1}. \tag{2.8}$$

Here, $e^{t\mathcal{Q}\mathcal{L}}$ will also be interpreted as a Koopman operator associated with the generator $\mathcal{Q}\mathcal{L}$. Since $\mathcal{Q}\mathcal{L} = \mathcal{L} - \mathcal{P}\mathcal{L}$, it can be written as an integro-differential operator.

A very important issue is the choice of ρ_0. A natural choice is an equilibrium probability density: $\rho_0 = \rho_{eq}$, where ρ_{eq} satisfies $\mathcal{L}^* \rho_{eq} = 0$. The most common choice is the canonical ensemble for ρ_{eq} (*e.g.*, by Mori [70]),

$$\rho_{eq} = \frac{1}{Z} e^{-\beta H}. \tag{2.9}$$

In this case, $a(\cdot,t)$ is simply a stationary random process with zero average. The GLE (2.7) is still useful since it describes the fluctuation of the CG variable. When the system is near equilibrium, this serves as the first approximation. Further corrections can be made using the linear response approach [95]. In general, the initial density ρ_0 can be constructed using the maximum entropy principle, as discussed in the monograph [110]. For example, given the averages of the local energy, this approach yields a probability density that is similar to the canonical ensemble (2.9) but with non-uniform temperature.

Several properties can be deduced from the derivation. Some of them have been discussed in the original work of Mori [71]. They are summarized as follows.

PROPOSITION 2.1. *Assuming that* $\langle a\rangle = 0$, *then the following properties hold,*

$$\begin{aligned} \langle F(\cdot,t)\rangle &= 0, \qquad && \forall t \geq 0, \\ \langle F(\cdot,t), a\rangle &= 0, && \forall t \geq 0. \\ \theta(t_1 - t_2) &= \langle F(\cdot,t_1), F(\cdot,t_2)\rangle M^{-1}, && \forall t_1, t_2 \geq 0 \text{ and } t_1 \geq t_2, \end{aligned} \tag{2.10}$$

Proof. Note that with the inner product defined above, the adjoint operator of $\mathcal{L}$ is $-\mathcal{L}$, and $\mathcal{P}$ and $\mathcal{Q}$ are self-adjoint, *i.e.*, $\langle \mathcal{L}b, c\rangle = -\langle b, \mathcal{L}c\rangle$ and $\langle \mathcal{P}b, c\rangle = \langle b, \mathcal{P}c\rangle$ for any $b, c \in \mathcal{H}$.

Now for the first property, we proceed as follows,

$$\begin{aligned} \langle F(\cdot,t)\rangle &= \langle \mathcal{Q}\mathcal{L}e^{t\mathcal{Q}\mathcal{L}}a\rangle = \langle \mathcal{L}e^{t\mathcal{Q}\mathcal{L}}a\rangle - \langle \mathcal{P}\mathcal{L}e^{t\mathcal{Q}\mathcal{L}}a\rangle \\ &= -\langle e^{t\mathcal{Q}\mathcal{L}}a\mathcal{L}1\rangle - \langle \mathcal{L}e^{t\mathcal{Q}\mathcal{L}}a, a\rangle M^{-1}\langle a\rangle = 0. \end{aligned} \tag{2.11}$$

For the second property, since $\mathcal{P}a = a$ and $\mathcal{Q}F(\cdot,t) = F(\cdot,t)$, one can easily verify that,

$$\langle F(\cdot,t), a\rangle = \langle \mathcal{Q}F(\cdot,t), \mathcal{P}a\rangle = \langle F(\cdot,t), \mathcal{Q}\mathcal{P}a\rangle = 0. \tag{2.12}$$

Finally, we start with the second equation in (2.8) and we get,

$$\begin{aligned} \theta(t_1 - t_2) &= -\langle \mathcal{L}F(\cdot, t_1 - t_2), a\rangle M^{-1} = \langle e^{(t_1 - t_2)\mathcal{Q}\mathcal{L}}\mathcal{Q}\mathcal{L}a, \mathcal{L}a\rangle M^{-1} \\ &= \langle \mathcal{Q}F(\cdot,t_1), e^{t_2\mathcal{L}\mathcal{Q}}\mathcal{L}a\rangle M^{-1} = \langle F(\cdot,t_1), \mathcal{Q}e^{t_2\mathcal{L}\mathcal{Q}}\mathcal{L}a\rangle M^{-1} \\ &= \langle F(\cdot,t_1), F(\cdot,t_2)\rangle M^{-1}. \end{aligned} \tag{2.13}$$

∎

The first and the third equations imply that the random process, $F(\cdot,t)$, is a stationary random process in the wide sense [18], and the relation (2.13) between the random noise and the memory kernel is known as the second fluctuation-dissipation theorem (FDT). It is a necessary condition for the solution to have the correct variance [54]. The second condition suggests that the random force and the initial value of a are uncorrelated. The last two properties have also been discussed in Mori's original paper [70].

Using the first property, one can take an average of the GLE (2.7), and arrive at a deterministic system,

$$\frac{d}{dt}\langle a\rangle(t) = \Omega\langle a\rangle(t) - \int_0^t \theta(t-s)\langle a\rangle(s)ds, \tag{2.14}$$

which is a set of integral-differential equations describing the time evolution of the average of $a(\cdot,t)$ [20]. This is often seen as a particular advantage of Mori's projection. However, in this paper, we will be more concerned with the quantities *with* fluctuations.

2.2. Approximations of kernel functions with Markovian embedding. A well known practical issue associated with the solution of the GLE is the computation of the memory term. Clearly, a direct evaluation of the integral requires the storage of the solutions from all previous steps, and such evaluations have to be carried out at every time step. A more delicate issue is that the kernel function $\theta(t)$,

$$\theta(t) = -\langle \mathcal{L}e^{t\mathcal{Q}\mathcal{L}}\mathcal{Q}\mathcal{L}a, a\rangle M^{-1}, \tag{2.15}$$

is determined by the orthogonal dynamics $e^{t\mathcal{Q}\mathcal{L}}$ for $t>0$. Except for very special cases, *e.g.*, linear dynamics, there is no explicit formula to compute $\theta(t)$ directly. Simulating the orthogonal dynamics is clearly not practical, either. To alleviate such effort, and more importantly, to make connections to existing generalized heat conduction models, we will use the Markovian embedded technique and approximate the memory term via an extended system of equations [56]. The idea is to incorporate the aforementioned values of the kernel function into the Laplace transform of θ. More specifically, we define

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

As $\lambda\to 0+$, using integration by parts repeatedly, we find that

$$\Theta(\lambda)=\lambda\theta(0)+\lambda^2\theta'(0)+\lambda^3\theta''(0)+\lambda^4\theta'''(0)+\cdots, \tag{2.17}$$

where derivatives of $\theta(0)$ are available as shown in the following. Recall that $\theta(t)=-\langle\mathcal{L}e^{t\mathcal{Q}\mathcal{L}}\mathcal{Q}\mathcal{L}a,a\rangle M^{-1}$. Direct computation yields,

$$\begin{aligned}
\theta(0)&=-\langle\mathcal{L}\mathcal{Q}\mathcal{L}a,a\rangle M^{-1},\\
\theta'(0)&=-\langle\mathcal{L}(\mathcal{Q}\mathcal{L})^2a,a\rangle M^{-1},\\
\theta''(0)&=-\langle\mathcal{L}(\mathcal{Q}\mathcal{L})^3a,a\rangle M^{-1},\\
&\cdots\cdots
\end{aligned} \tag{2.18}$$

Let M_j be the *normalized moments* associated with the statistics of a, *i.e.*,

$$M_j:=\langle\mathcal{L}^ja,a\rangle M^{-1}. \tag{2.19}$$

Using the notation of M_j, we are able to further simplify the above derivatives. The first few terms are listed as follows,

$$\begin{aligned}
\theta(0)&=-M_2+M_1^2,\\
\theta'(0)&=-M_3+M_2M_1+M_1M_2-M_1^3,\\
\theta''(0)&=-M_4+M_3M_1+M_2^2+M_1M_3-M_2M_1^2-M_1M_2M_1-M_1^2M_2+M_1^4,\\
&\cdots\cdots
\end{aligned} \tag{2.20}$$

Given statistics of initial states, the above quantities can be accurately calculated.

We now turn to the limit as $\lambda\to+\infty$, which embodies long-time behavior of the kernel function. For this calculation, we start with the GLE (2.7), multiplying both sides by $a^\intercal$ and taking the average. Let $M(t)=\langle a(y_0,t),a\rangle$, which represents two-point statistics, then we have,

$$\dot{M}(t)=\Omega M(t)-\int_0^t\theta(t-s)M(s)ds. \tag{2.21}$$

Let $\widetilde{M}(\lambda)$ be the Laplace transform of $M(t)$. Taking the Laplace transform of (2.21), we find,

$$\frac{1}{\lambda}\widetilde{M}(\lambda)-M(0)=\Omega\widetilde{M}(\lambda)-\Theta(\lambda)\widetilde{M}(\lambda), \tag{2.22}$$

which yields,

$$\Theta(+\infty)=\Omega+MC^{-1},\ M=M(0),\ C:=\lim_{\lambda\to+\infty}\int_0^{+\infty}e^{-t/\lambda}M(t)dt. \tag{2.23}$$

In practice, we keep the exponential penalty from the Laplace transform to ensure the integral is finite. It is well known that [77] a necessary condition for the integral to be convergent is that the process $a(\cdot,t)$ is ergodic, but it is not sufficient. Finally, for a scalar CG variable, the second term can be identified as the the correlation length.

One can see that $\Theta(+\infty)$ is again related to the statistics of $a(\cdot,t)$, which can be obtained either from the full model or from experimental observations. We now incorporate the values of Θ from both regimes: $\lambda\to 0+$ and $\lambda\to+\infty$. Such two-sided approximations, which are similar to the Hermite interpolation problems, have demonstrated promising accuracy over both short and long-time scales [56].

The idea of the Markovian embedding is to approximate the memory term by rational functions in terms of the Laplace transform. In general, we can consider a rational function in the following form,

$$R_{k,k}=\left[I-\lambda B_1-\cdots-\lambda^k B_k\right]^{-1}\left[A_0+\lambda A_1+\lambda^2 A_2+\cdots+\lambda^k A_k\right]. \tag{2.24}$$

The coefficients in the rational function can be determined based on the values of the kernel functions, *e.g.*, those presented in the previous section.

When $k=0$, we are led to a constant function, $R_{0,0}=\Gamma$, which, we choose to be given by (2.23):

$$\Gamma=\Theta(+\infty). \tag{2.25}$$

In the time domain, this amounts to approximating the kernel function by a delta function:

$$\int_0^t\theta(t-s)a(\cdot,s)ds\approx\Gamma a(\cdot,t). \tag{2.26}$$

This is often referred to as the Markovian approximation [43, 50].

When $k=1$, we have,

$$R_{1,1}(\lambda)=\left[I-\lambda B_1\right]^{-1}\left[A_0+\lambda A_1\right]. \tag{2.27}$$

To determine the coefficients, we match the following values,

$$R_{1,1}(0_+)=\Theta(0_+),\ R'_{1,1}(0_+)=\Theta'(0_+)\ \text{ and }\ R_{1,1}(+\infty)=\Theta(+\infty), \tag{2.28}$$

which yield,

$$A_0=0,\ A_1=\theta(0)\ \text{ and }\ B_1=-A_1\Theta(+\infty)^{-1}. \tag{2.29}$$

In the time-domain, this corresponds to an approximation of the kernel function by an matrix exponential $e^{B_1 t}A_1$. At this point, if we define the memory term as $z(\cdot,t)$,

$$z(\cdot,t)=\int_0^t\theta(t-s)a(\cdot,s)ds, \tag{2.30}$$

we can write down an auxiliary equation,

$$\dot{z}(\cdot,t)=A_1a(\cdot,t)+B_1z(\cdot,t). \tag{2.31}$$

This way, the memory term is embedded in an extended dynamical system *without* memory.

REMARK 2.2. One may notice that the stochastic equations derived from this scheme form a closed system, without any explicit dependence on the initial condition y_0. Therefore, we will simply write $a(y_0,t)$ as $a(t)$ in all the following reduced models.

As the order of the approximation k increases, we obtain a hierarchy of approximations for the memory term, which can be written as a larger extended system of equations [56,66]. We will not discuss the higher order approximations in this paper.

REMARK 2.3. The idea of introducing auxiliary variables and extended dynamics has been implemented in various previous works, *e.g.*, [6,12,24,34,63,64]. As alluded to in (2.15), we notice again that in general, the memory kernel function in the time-domain is not accessible. Therefore, a direct interpolation by a sum of exponential functions, *e.g.*, [6,34], is not feasible. Our formulation rests upon a rational interpolation, in which the required function values are statistics of the coarse-grain variables.

It remains to approximate the random noise term. This will be discussed in the next section, along with a specific example of the MD model.

3. Application to energy transport

We now turn to the modeling of energy transport problem, especially the heat conduction process. We will explain how the CG variables are defined using a one-dimensional example, and then present the resulting reduced models. Our emphasis will be placed on the modeling of the random noise.

3.1. A one-dimensional example. Let's consider a one-dimensional isolated chain model of N atoms and they are evenly divided into n blocks, each of which contains ℓ atoms, as shown in Figure 3.1; $N=n\ell$. The spacing of two atoms is ε_0 at equilibrium. We will focus on the study of energy transport between these blocks.

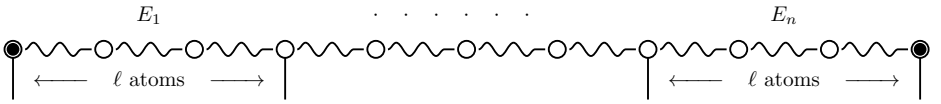

FIG. 3.1. *1-D chain of atoms. Every ℓ atoms are grouped into one block.*

Let x and v be the displacements and velocities respectively, satisfying (2.1). Let S_I be the index set of I-th block, labeled as, $S_I=\{\ell(I-1)+1,\cdots,\ell I\}$. Let $\phi(x_i-x_j)$ be the pairwise potential coming from interactions between the ith and jth atoms. Here, we use the Fermi-Pasta-Ulam (FPU) potential in the form of,

$$\phi(r)=\frac{r^2}{2}+c_0\frac{r^3}{3}+c_1\frac{r^4}{4}, \tag{3.1}$$

where c_0,c_1 are parameters in the potential. We assume $c_1>0$. If we only consider the nearest neighbor interactions, the potential energy of this 1-d chain is given by,

$$V(x)=\sum_{i=1}^{N}\phi(x_i-x_{i-1}), \tag{3.2}$$

where the periodic boundary condition is imposed, *i.e.*, $x_0=x_N$ and $x_{N+1}=x_1$.

We define the locally averaged energy associated with the I-th block as follows,

$$E_I = \frac{1}{\ell\varepsilon_0}\sum_{i\in S_I}\frac{1}{2}mv_i^2+\frac{1}{2}\phi(x_i-x_{i-1})+\frac{1}{2}\phi(x_{i+1}-x_i). \tag{3.3}$$

The rate of change of the local energy can be attributed to the energy flux J,

$$\dot{E}_I = -\frac{1}{\ell\varepsilon_0}\Big(J_{I+\frac{1}{2}}-J_{I-\frac{1}{2}}\Big), \tag{3.4}$$

where $J_{I+\frac{1}{2}}$ is the energy flux between the Ith and $(I+1)$th blocks. Direct calculation yields,

$$J_{I+\frac{1}{2}} = -\frac{1}{2}\phi'(x_{\ell I+1}-x_{\ell I})(v_{\ell I+1}+v_{\ell I}). \tag{3.5}$$

Notice that the energy flux only depends on the atoms next to the interfaces between two adjacent blocks. This expression has also been used in many numerical studies, *e.g.*, in [55, 100, 105].

REMARK 3.1. The energy balance, in principle, is exact, but the equation is not closed. The Mori-Zwanzig formalism provides a closure model with noise that takes into account the under-resolved dynamics.

As an application of the Mori's projection method, we define the CG variables as the centered averaged energy of blocks,

$$a_I(\cdot,t) := E_I(\cdot,t)-\langle E_I(\cdot,0)\rangle. \tag{3.6}$$

The subtraction of the average is to ensure that $\langle a\rangle = 0$.

For the initial configuration, we assume that the local energy is given, denoted here by $\bar{E}$. One approach to set up the initial density ρ_0 is to maximize the entropy $-\int\rho\log\rho dy_0$, subject to the constraints, $\langle E_I\rangle = \bar{E}_I$. This approach yields the following density function [62],

$$\rho_0 \propto \exp-\sum_I \beta_I E_I, \tag{3.7}$$

with β_I being the Lagrange multiplier. It plays the role of local inverse temperature, and it can be determined based on the constraints $\langle E_I\rangle = \bar{E}_I$.

With this choice of ρ_0, one can show that in the GLE (2.7) $\Omega\equiv 0$. So it is enough to approximate the memory term. In the approximation of the memory term, we will approximate the density $\rho_0 \approx \frac{1}{Z}\exp-\beta H$, with β being a reference temperature. The approximation considerably simplifies the calculations. Let us denote $\bar{E} := \langle E(\cdot,0)\rangle$. Direct computation yields, for $j\geq 0$,

$$\begin{aligned} M_{2j+1} &= \langle \mathcal{L}^{2j+1}a, a\rangle M^{-1} \\ &= \langle \mathcal{L}^{2j+1}E(\cdot,0), E(\cdot,0)-\bar{E}\rangle M^{-1} - \langle \mathcal{L}^{2j+1}\bar{E}, E(\cdot,0)-\bar{E}\rangle M^{-1} \\ &= \langle \mathcal{L}^{2j+1}E(\cdot,0), E(\cdot,0)\rangle M^{-1}. \end{aligned} \tag{3.8}$$

From simple algebraic observations, $\mathcal{L}^{2j+1}E(y_0,0)$ is odd w.r.t. v_0 and $E(y_0,0)$ is even w.r.t. v_0. We integrate the product over the velocity domain weighted by a Gaussian

distribution. Due to the symmetry, the integral is zero. This implies that the odd moments of a vanish in (2.20), *i.e.*,

$$M_{2j+1}=0, \quad \forall j\geq 0. \tag{3.9}$$

With the above results, we are able to further simplify the formulas (2.20) for the derivatives of θ at $t=0$,

$$\begin{aligned}
&\theta(0)=-M_2,\\
&\theta'(0)=0,\\
&\theta''(0)=-M_4+M_2^2,\\
&\theta'''(0)=0,\\
&\theta''''(0)=-M_6+M_4M_2+M_2M_4+M_2^3,\\
&\cdots\cdots
\end{aligned} \tag{3.10}$$

These quantities will be used later as interpolation conditions to determine the kernel function.

3.2. Deterministic models. Before we discuss the approximation of the random noise term, let us first briefly discuss the deterministic GLE model (2.14). In particular, we will examine the models obtained from various approximations of the kernel function in the memory term.

For the zeroth order approximation (2.26), the memory effect is eliminated entirely, with the kernel function approximated by,

$$\theta(t)\approx\Gamma\delta(t), \tag{3.11}$$

where Γ is given by the formulas (2.25) and (2.23). In this case, the GLE (2.14) is reduced to an ODE system,

$$\frac{d}{dt}\langle a\rangle(t)=-\Gamma\langle a\rangle(t), \quad \langle a\rangle(0)=\bar{E}. \tag{3.12}$$

For our energy transport problem (3.3), the matrix is given by,

$$\Gamma=MC^{-1}, \tag{3.13}$$

with M,C defined in (2.23). In light of the periodic boundary conditions and the fact that the system is partitioned uniformly to define the local energy, it is reasonable to assume that a is a stationary process in space, implying that M and C are both circulant matrices [25]. They can be generated from just one row of the matrices. This also implies that Γ is a circulant matrix as well.

We have observed from various numerical tests that $-\Gamma$ is proportional to a discrete Laplacian operator,

$$\Gamma\propto-\nabla_h^2, \quad h=\ell\varepsilon_0, \tag{3.14}$$

as suggested in the left figure in Figure 3.2. With the approximation that $\Gamma=-\alpha\nabla_h^2$, Equation (3.12) coincides with a direct discretization of the one-dimensional heat equation,

$$\frac{d}{dt}\langle a\rangle(t)=\alpha\nabla^2\langle a\rangle(t). \tag{3.15}$$

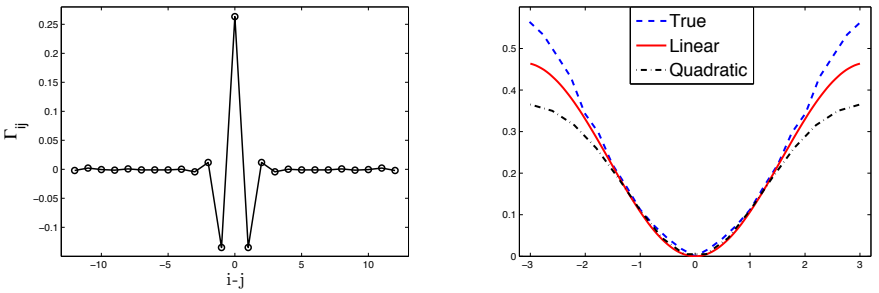

FIG. 3.2. *The left figure shows the elements of* Γ*, with* x*-axis being indices* $i-j$ *and* y*-axis being* Γ_{ij}*. The right figure shows the true value of* $\widehat{\gamma}$ *and its linear and quadratic fitting of Laplace operator in Fourier space.*

The parameter α represents the diffusion coefficient of the energy. To obtain the actual value of the parameter, we may expand Γ in terms of the finite-difference operator ∇_h^2:

$$-\Gamma \approx \alpha \nabla_h^2 + \nu \nabla_h^4 + \cdots. \tag{3.16}$$

A more transparent calculation, however, can be done using discrete Fourier transforms, thanks to the circulant structures. More specifically, a discrete Fourier transform can be applied to the generator of a circulant matrix [25]. It will diagonalize the matrix and yield its eigenvalues. We let $\widehat{\gamma}(k)$ and $\widehat{\lambda}(k)$ be the corresponding Fourier transforms of Γ and the discrete Laplacian operator, $k \in (-\frac{\pi}{h}, \frac{\pi}{h}]$. The wave numbers are uniformly distributed in the interval $(-\frac{\pi}{h}, \frac{\pi}{h}]$, and they can be viewed as k-points. In particular, $\widehat{\lambda}(k) = \frac{1}{h^2}[2-2\cos(k)]$. Similarly, we let $\widehat{m}(k)$ and $\widehat{c}(k)$ be the Fourier transform of the two matrices M and C, respectively; then the matrix multiplications are turned into convolution, and we have,

$$\widehat{\gamma}(k) = \widehat{m}(k)/\widehat{c}(k). \tag{3.17}$$

The approximation in (3.16) corresponds to an expansion of $-\widehat{\gamma}(k)$ around $k=0$,

$$\widehat{\gamma}(k) = \alpha\widehat{\lambda}(k) + \nu\widehat{\lambda}(k)^2 + \mathcal{O}\big(\widehat{\lambda}(k)^4\big), \quad |k| \ll 1, \tag{3.18}$$

as shown in Figure 3.2. In particular, we can choose α as

$$\alpha = \lim_{k\to 0} \frac{\widehat{\gamma}(k)}{\widehat{\lambda}(k)}, \tag{3.19}$$

where the limit simply indicates a selection of a small value of k, *e.g.*, the k-point that is closest to the origin. We can not, however, choose $k=0$, since both terms are zero at that point.

From the numerical observations, M is close to being proportional to an identity matrix, and $\widehat{m}(k)$ is a constant that does not depend on k. It suffices to check $\widehat{c}(k)$ for

small k. Direct computation yields,

$$\begin{aligned}\widehat{c}(k)\widehat{\lambda}(k) &= \frac{1}{n}\int_0^\infty \langle \widehat{a}(k,t), \widehat{a}(k,0)^*\rangle dt \frac{\left(e^{ik}-1\right)\left(e^{-ik}-1\right)}{h^2} \\ &= \frac{1}{n}\int_0^\infty \langle \frac{\left(e^{ik}-1\right)\widehat{a}(k,t)}{h}, \frac{\left(e^{-ik}-1\right)\widehat{a}(k,0)^*}{h}\rangle dt, \quad 0<|k|\ll 1 \\ &= \frac{1}{n}\int_0^\infty \langle \widehat{\nabla_h a}(k,t), \widehat{\nabla_h a}(k,0)^*\rangle dt.\end{aligned} \tag{3.20}$$

Interestingly, the right-hand side corresponds to the Fourier transform of the gradient of a. In fact, it is the Fourier transform of the covariance,

$$D := \int_0^{+\infty} \langle \nabla_h a(\cdot,t), \nabla_h a(\cdot,0)\rangle dt. \tag{3.21}$$

By combining (3.17), (3.19) and (3.20), we could determine the parameter α as

$$\alpha = \lim_{k\to 0} \frac{\widehat{m}(k)}{\widehat{d}(k)}, \tag{3.22}$$

where $\widehat{d}$ is the Fourier transform of D (3.21). It is clearly possible to include the $\widehat{\lambda}^4$ term in (3.17) and derive a higher order model. But that will not be pursued here.

By a linearization of the average energy with respect to the temperature, $\langle a\rangle = c_p T$, with c_p being the specific heat, one can substitute the averaged energy by $c_p T$ in (3.15) and obtain

$$\frac{d}{dt}\langle a\rangle = \alpha c_p \nabla^2 T = -\nabla\cdot(-\alpha c_p \nabla T). \tag{3.23}$$

Combining the above equation with the conservation of energy, $\frac{d}{dt}\langle a\rangle = -\nabla\cdot\langle j\rangle$, one has

$$\langle j\rangle = -\alpha c_p \nabla T, \tag{3.24}$$

which implies that the thermal conductivity $\kappa_{\mathrm{CG}} = \alpha c_p$. The heat capacity c_p can be determined from the statistics of the local energy [62]: $c_p = h\mathrm{var}(E_I)/k_B T^2$, where h appears due to the fact that E_I is averaged over blocks (energy per unit volume). This leads to a formula for the thermal conductivity,

$$\kappa_{\mathrm{CG}} = \lim_{k\to 0} h\mathrm{var}(E_I)\widehat{m}(k)/(\widehat{d}(k)k_B T^2). \tag{3.25}$$

The expression on the right-hand side corresponds to the Fourier transform of the matrix

$$\frac{h}{k_B T^2}\langle a,a\rangle^2\left[\int_0^{+\infty}\langle \nabla_h a(\cdot,t), \nabla_h a(\cdot,0)\rangle dt\right]^{-1}. \tag{3.26}$$

Meanwhile, there are two conventional approaches to determine the heat conductivity. The first approach is based on the Green-Kubo formula, which is based on a linear response theory. The coefficient can be determined from equilibrium molecular

dynamics simulations, which has been widely implemented [8, 58, 59, 88]. The formula is written in terms of the autocorrelation function of the heat current,

$$\kappa_{\mathrm{GK}} = \frac{1}{k_B T^2 N \varepsilon_0} \int_0^\infty \langle \mathbf{J}(t)\mathbf{J}(0)\rangle \, dt, \tag{3.27}$$

where $\mathbf{J} = \varepsilon_0 \sum_{I=1}^{n} J_{I+\frac{1}{2}}$.

The other traditional approach is based on non-equilibrium MD simulations with two different temperatures maintained at the boundary. The thermal conductivity can be estimated from the heat flux induced by a temperature gradient,

$$\langle j \rangle = -\kappa_{\mathrm{NEMD}} \nabla T. \tag{3.28}$$

We conducted several numerical tests and calculated the heat conductivity based on the three formulas. In our calculation, we consider 250 atoms, divided into 25 blocks equally. The mass of each atom is set to unity $m = 1$, and the spacing of two atoms is set to $\varepsilon_0 = 1$. For equilibrium MD simulations, we prepare the system with two Nosé-Hoover thermostats at temperature 1.5 to reach steady states and use periodic boundary conditions afterwards. For non-equilibrium simulations, we impose the two Nosé-Hoover thermostats at temperatures of 1.4 and 1.6, at the left and right boundaries, respectively. The left- and right-most particles in the chain are fixed. For the integration in time, the 6-th order symplectic method is employed [106]. Each MD simulation run consists of 10^7 time steps with stepsize $\Delta t = 0.02$. We sample the data every 5 steps. Motivated by the studies in [92], we tested both the symmetric $(c_0 = 0, c_1 = 1)$ and the asymmetric $(c_0 = 1, c_1 = 1)$ FPU potentials. The results are listed in Table 3.1. The results show qualitative agreement.

In addition, the same MD system is used for all the subsequent numerical simulations as in this section.

thermal conductivity	CG models	Green-Kubo	NEMD
symmetric potential	31.82 ± 1.41	30.18 ± 1.34	29.16 ± 0.97
asymmetric potential	35.78 ± 1.40	34.39 ± 1.47	35.17 ± 0.89

TABLE 3.1. *This table shows the numerical results of heat conductivity computed with three different methods. In the Green-Kubo formula, to truncate the integral, we have multiplied the correlation function with an exponential decay penalty* $e^{-t/\lambda}$ *with* $\lambda = 10^3$ *to eliminate the contribution from the long-time correlation, which may be subject to large numerical and sampling error. For NEMD, we run* 5 *copies and use the average to compute the thermal conductivity. For each method, we repeat the experiments* 10 *times to obtain the error bar.*

We now examine the model obtained from the first order approximation (2.31) and compare it to existing continuum models. With the rational approximation and the introduction of the auxiliary variable, we have,

$$\begin{cases} \dot{\langle a \rangle}(t) = -z(t), \\ \dot{z}(t) = B_1 z(t) + A_1 \langle a \rangle(t), \end{cases} \tag{3.29}$$

where A_1 and B_1 are determined in (2.29). The noise has been averaged out. The second interpolation condition in (2.29) yields $A_1 = -M_2$, which according to (3.4) and (3.8), is proportional to a discrete Laplacian operator: $A_1 \propto \nabla_h^2$. Meanwhile, the third

interpolation condition in (2.29) leads us to $B_1 = -A_1\Gamma^{-1}$. From the previous discussion, $\Gamma \approx -\alpha\nabla_h^2$. As a result, we have

$$A_1 = -c^2\nabla_h^2, \quad B_1 \approx \frac{c^2}{\alpha} I. \tag{3.30}$$

Altogether, this leads to a second order equation

$$\ddot{\langle a\rangle}(t) = -\frac{c^2}{\alpha}\dot{\langle a\rangle}(t) + c^2\nabla_h^2\langle a\rangle(t). \tag{3.31}$$

Interestingly, this coincides with the Cattaneo and Vernotte (CV) model in heat transport, explicitly demonstrated in [11,98], where c can be interpreted as the second sound speed and is determined from the relation. Higher order approximations will correspond to higher order relaxation models, similar to the models by Tzou [96].

The two cases presented here illustrate how the memory term can be approximated using the rational approximation in terms of the Laplace transform. In particular, it gives rise to deterministic (or drift) terms in the resulting approximate models. However, in this paper, we will mainly focus on the fluctuating effects in energy transport models. By introducing additive and multiplicative noises to the existing models, one obtains a set of SDEs, and in the next section, we will examine the resulting statistics.

3.3. Approximations of the noise. Before we discuss the approximation of the random noise, we first start with an equilibrium MD simulation, in which we observe the trajectories of energy $a(\cdot,t)$ and noise term $F(\cdot,t)$ at the steady state. Since the initial values do not play an explicit role in the behaviors at steady states, we denote them as $a(t)$ and $F(t)$ for simplicity.

The same MD system as in Section 3.2 is used here and for all subsequent numerical experiments. The histograms of one entry (the 15th block) of a and F at equilibrium are shown in Figure 3.3, which can be regarded as the true stationary distributions. (Here it is clear that $a(t)$ is stationary. We will also assume that $F(t)$ is stationary as well.) We use $F(0) = La$ to generate the histogram of $F(t)$ due to the stationarity of the process. Interestingly, both quantities exhibit non-Gaussian statistics. The PDF of $a(t)$ fits perfectly to a shifted Gamma distribution, and the PDF of the random noise follows a Laplace distribution. Furthermore, we observe that the locally averaged energy obeys the Gamma distribution when the potential is either harmonic or asymmetric, as shown in Figure 3.4.

The emergence of the non-Gaussian statistics can be understood heuristically as follows. In the harmonic case, where the interactions are modelled by springs, we have the total energy given by,

$$E = \sum_i \frac{\sigma}{2}(x_{i+1} - x_i)^2 + \frac{v_i^2}{2}. \tag{3.32}$$

The parameter σ is the spring constant. Therefore the relative displacement $x_{i+1} - x_i$ can be viewed as independent normal random variables with variance k_BT/σ.

The local energy within each block can thus be written as,

$$E_I = \frac{1}{4\ell\varepsilon_0}\left[d_{I\ell}^2 - d_{(I+1)\ell}^2\right] + \frac{1}{2\ell\varepsilon_0}\sum_{i\in S_I}\left[d_i^2 + v_i^2\right], \quad d_i = \sqrt{\sigma}(x_{i+1} - x_i). \tag{3.33}$$

For large ℓ, the variable E_I will be mostly determined by the summation which is a Gamma distribution with parameters related to the block size.

On the other hand, with the harmonic approximation the noise term becomes,

$$F_I = \mathcal{L}a = -\frac{1}{h}\left(J_{I+\frac{1}{2}} - J_{I-\frac{1}{2}}\right), \tag{3.34}$$

where

$$J_{I+\frac{1}{2}} = -\frac{\sigma}{2}(x_{\ell I+1} - x_{\ell I})(v_{\ell I+1} + v_{\ell I}) \tag{3.35}$$

from (3.5). Hence, it can be written as $F = \xi_1\xi_2 - \xi_3\xi_4$ where ξ_1, ξ_2, ξ_3, and ξ_4 are i.i.d. normal random variables, which according to the well known result [53], leads to the Laplace distribution.

This explains heuristically that when the interactions are nearly harmonic, such statistics are expected. However, the presence of anharmonic terms will affect the statistics, and it leads to different Gamma or Laplace distributions, as shown in Figure 3.3.

In the following section, we will focus on the approximation of the noise term so that the solution of the reduced models gives consistent statistics of the local energy. Here, we will consider both additive and multiplicative noises.

3.3.1. Approximation by additive noise. A natural (and most widely used) approximation is by a Gaussian white noise. For instance, for the first (Markovian) approximation (2.23), we are led to a linear SDE,

$$\dot{a}(t) = -\Gamma a(t) + \sigma\zeta(t), \tag{3.36}$$

where $\zeta(t)$ is the standard Gaussian-white noise,

$$\langle \zeta_i(t_1), \zeta_j(t_2)\rangle = \delta_{ij}\delta(t_1 - t_2). \tag{3.37}$$

In order for the solution a to have the correct covariance $M = \langle a, a\rangle$, the parameter σ has to satisfy the Lyapunov equation [77, 83],

$$\Sigma := \sigma\sigma^{\mathsf{T}} = \Gamma M + M\Gamma^{\mathsf{T}}. \tag{3.38}$$

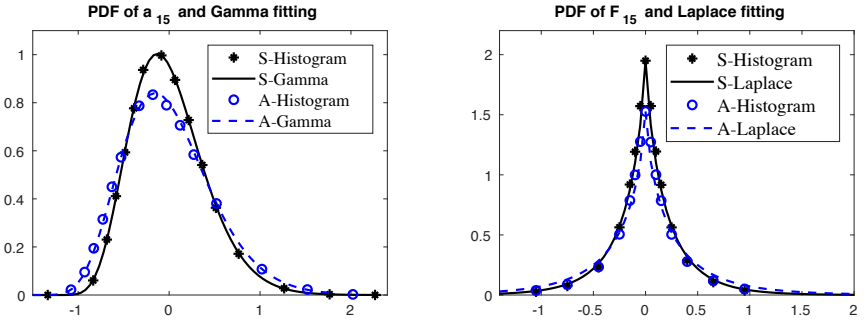

FIG. 3.3. *The left figure shows the histogram of a_{15} at equilibrium with symmetric potential ($c_0 = 0, c_1 = 1$) and asymmetric potential ($c_0 = 1, c_1 = 1$). The data fit well to shifted Gamma distributions $\Gamma_{k,\eta}(x) \sim (x+\mu)^{k-1}\exp(-\eta(x+\mu))$, $x \geq -\mu$, with parameters $k = 10.3816$, $\eta = 7.7752$ and $\mu = 1.3352$ and $k = 10.2011$, $\eta = 6.4414$ and $\mu = 1.5837$, for symmetric and asymmetric potentials respectively. The right figure shows the histogram of F_{15} at equilibrium with symmetric and asymmetric potentials, and their fitting to Laplace-type distributions $Lap_\gamma(x) \sim \exp(-\gamma|x|)$ with parameters $\gamma = 3.8986$ and $\gamma = 3.0538$.*

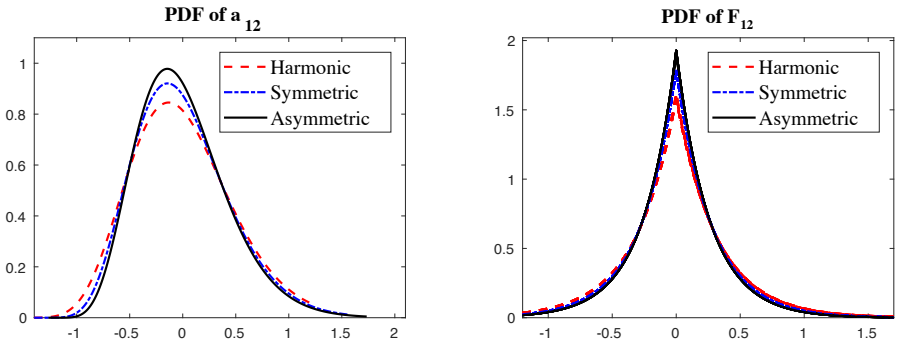

FIG. 3.4. *These two figures show the histograms of a_{12} and F_{12} at equilibrium from MD simulations, governed by different potential energies. The potentials are given by $\phi(r)=\frac{r^2}{2}+c_0\frac{r^3}{3}+c_1\frac{r^4}{4}$,, with harmonic label $c_0=0,c_1=0$, symmetric label $c_0=0,c_1=1$ and asymmetric $c_0=1,c_1=1$. The block size and temperature are the same for three cases in MD simulations, which accounts for the similarity of the parameters.*

For the next order rational approximation (2.28) of the kernel function, we may introduce noise via the second equation. Namely,

$$\begin{cases} \dot{a}(t) = -z(t), \\ \dot{z}(t) = B_1 z(t) + A_1 a(t) + \sigma\zeta(t), \end{cases} \tag{3.39}$$

where A_1 and B_1 are determined in (2.29). As demonstrated in the previous section, the deterministic part of the model coincides with the CV model. Here we discuss how the noise can be introduced to guarantee the correct statistics.

In (3.39), it is clear that the second equation can be solved explicitly and substituted into the first equation, which would yield a similar equation to the GLE (2.7). By choosing the initial condition $z(0)$ and Σ appropriately, the approximations to the memory kernel and the random noise can be made consistent, in terms of the second FDT (2.10).

THEOREM 3.1. *Assuming the covariance of $z(0)$ is A_1, and*

$$B_1 A_1 + A_1 B_1^{\intercal} + \Sigma = 0, \tag{3.40}$$

then, the extended system is equivalent to approximating the kernel function by $\theta_1(t) = e^{tB_1}A_1$, and the approximate noise, denoted by $F_1(t)$, to $F(t)$ satisfies the second FDT exactly. Namely,

$$\theta_1(t-t') = \langle F_1(t), F_1(t')\rangle M^{-1}, \quad \forall\, t \geq t' \geq 0.$$

The proof of this theorem can be found in [66]. In light of (2.29) and the FDT, the matrix A_1 is semi-positive definite, and it can be used as a covariance matrix.

REMARK 3.2. The model (3.39) is a Langevin dynamics (by changing z to $-z$), which typically comes from the force (momentum) balance. It is therefore interesting that such models also arise in the balance equations for the energy and energy flux.

The approximation by additive noises inevitably leads to a Gaussian distribution for $a(t)$ [83]. To check the validity of this assumption, we solve the reduced models

(3.36) and (3.39) directly and compare the results with true statistics obtained from MD simulations, including histograms of solutions at steady states and autocorrelation functions.

From Figure 3.5, we observe that the autocorrelation of energy can be well captured from models driven by additive Gaussian noise, while the PDFs deviate from the true distributions. One obvious violation is that the solution of (3.36) is always Gaussian, which does not have a lower bound for energy and also unable to recover a non-symmetric PDF.

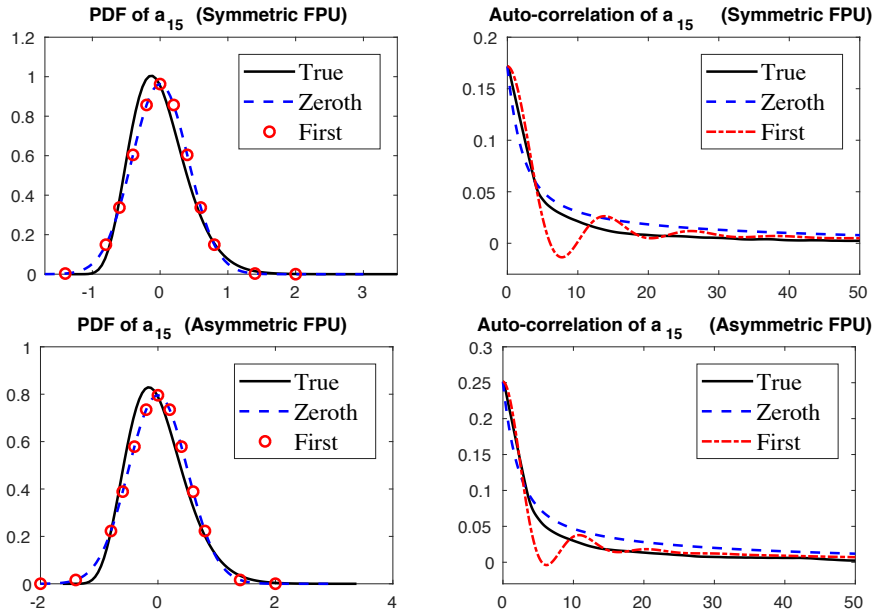

FIG. 3.5. *The left two figures show PDFs of a_{15} obtained from approximated models (3.36) and (3.39). The right figures show the auto time correlations $\langle a_{15}(t)a_{15}(0)\rangle$. Numerical results are compared with true statistics that come from direct MD simulations using a 6th order symplectic method.*

3.3.2. Approximation by Gaussian multiplicative noise. As alluded to in the previous section, the approximate model driven by Gaussian additive white noise may not capture the correct PDF. In this section, we consider the multiplicative noise, with the objective of enforcing the correct equilibrium statistics for the solution of the SDEs.

We start with a further observation that the energy of each block is almost independent to each other. Very interestingly, the same observations have been made for biomolecules [32]. Clearly, it is difficult to prove the independence theoretically as atoms in block boundaries share pair potentials with their neighbors in adjacent blocks and precisely the small correlation between blocks generates energy transport. However, when the block size is relatively large, the correlation is quite weak. Therefore, we keep this as our main assumption, and postulate the stationary PDF ($\rho(a)$) of the energy as a shifted multi-Gamma distribution with parameters k and η,

$$\rho(a) = \frac{1}{Z}\prod_{i=1}^{n}\left(a_i + \frac{k_i}{\eta_i}\right)^{k_i-1} e^{-\eta_i(a_i+k_i/\eta_i)},\ a_i \geq -\frac{k_i}{\eta_i}, \tag{3.41}$$

where

$$Z=\prod_{i=1}^{n} Z_i, \quad Z_i=\int_{-\frac{k_i}{\eta_i}}^{\infty}\left(a_i+\frac{k_i}{\eta_i}\right)^{k_i-1} e^{-\eta_i(a_i+k_i/\eta_i)}\, da_i. \tag{3.42}$$

Now we reconsider the zeroth order approximation, the Markovian approximation by $R_{0,0}=\Theta(+\infty)=\Gamma$. With a multiplicative noise, we are solving the following SDE,

$$\dot{a}(t)=-\Gamma a(t)+\sigma(a(t))\zeta(t), \tag{3.43}$$

where $\zeta(t)$ is again the standard Gaussian white noise. The SDE is interpreted in the Itô sense.

To derive a simple formula, we seek σ in a diagonal form. We aim to construct $D:=2\sigma\sigma^{\intercal}$ to ensure the desired PDF given by (3.41). By simplifying the Fokker-Planck equation (FPE) that corresponds to (3.43), we obtain,

$$\frac{\partial D_{ii}\rho}{\partial a_i}=-\rho\left(\Gamma_{ii}a_i+\sum_{j=1,j\neq i}^{n}\Gamma_{ij}a_j\right). \tag{3.44}$$

By directly solving these differential equations, we obtained an explicit formula for the matrix D.

THEOREM 3.2. *If Γ has only non-positive off-diagonal entries and Γ is semi-positive definite, then there exists a diagonal matrix D for which the multivariate Gamma distribution (3.41) is a steady state solution of the Fokker-Planck equation. The diagonals of D are given by a positive expression,*

$$D_{ii}=\frac{\Gamma_{ii}}{\eta_i}\left(a_i+\frac{k_i}{\eta_i}\right)-\sum_{j=1,j\neq i}^{m}\Gamma_{ij}\frac{a_j\int_{-\frac{k_i}{\eta_i}}^{a_i}\rho_i(x)dx+k_j/\eta_j}{\rho_i(a_i)}, \tag{3.45}$$

where ρ_i is the marginal PDF of a_i, $\rho_i=\frac{1}{Z_i}\left(a_i+\frac{k_i}{\eta_i}\right)^{k_i-1}e^{-\eta_i(a_i+k_i/\eta_i)}$.

Proof. Direct computation yields that the stationary PDF given in (3.41) satisfies the FPE of the first order model in (3.43), which proves this theorem. ∎

Let us turn to the model obtained by the first order approximation of the memory term. With multiplicative Gaussian white noise, the first order model can be written formally as follows,

$$\begin{cases}\dot{a}(t)=-z(t),\\ \dot{z}(t)=Aa(t)+Bz(t)+\sigma\big(a(t),z(t)\big)\zeta(t),\end{cases} \tag{3.46}$$

where $A=A_1$ and $B=B_1$ are given in (2.29). This is a Langevin equation, where we allow the diffusion coefficient to depend on both a and z.

Similar to the previous case, we are seeking an equilibrium distribution of a and z. Notice that $z(0)=F(0)$, which follows a Laplace distribution. We further assume that a and z are independent. These assumptions lead to the following ansatz for the joint PDF.

$$\rho(a,z)=\frac{1}{Z}\exp\Big[-\sum_{i=1}^{n}\big(W_i(a_i)+\gamma_i|z_i|\big)\Big]. \tag{3.47}$$

In order to find a reasonable expression of σ, we force σ to be diagonal and work with the steady state solution of the FPE, which writes as

$$\sum_{i=1}^{n}\frac{\partial}{\partial z_i}\left(\sum_{j=1}^{n}A_{ij}a_j\rho+\sum_{j=1}^{n}B_{ij}z_j\rho\right)=\sum_{i=1}^{n}\frac{\partial^2 D_{ii}\rho}{\partial z_i^2}, \tag{3.48}$$

where $D=2\sigma^2$. Furthermore, if it satisfies

$$\frac{\partial D_{ii}\rho}{\partial z_i}=\sum_{j=1}^{n}A_{ij}a_j\rho+\sum_{j=1}^{n}B_{ij}z_j\rho, \tag{3.49}$$

then ρ is a solution to the FPE in (3.48).

THEOREM 3.3. *Suppose diagonal entries of B are non-positive. ρ given in (3.47) is an equilibrium density of the FPE of first order model (3.46) if σ is a diagonal matrix and $\sigma_{ii}=(D_{ii}/2)^{1/2}$ for $i=1,2,\cdots,n$, where*

$$D_{ii}=-\frac{sgn(z_i)}{\gamma_i}\left(\sum_{j=1}^{n}A_{ij}a_j+\sum_{j=1,j\neq i}^{n}B_{ij}z_j\right)-B_{ii}\left(\frac{|z_i|}{\gamma_i}+\frac{1}{\gamma_i^2}\right)+\mathcal{D}_i e^{\gamma|z_i|}, \tag{3.50}$$

and

$$\mathcal{D}_i=\frac{1}{\gamma_i}\left(\sum_{j=1}^{n}|A_{ij}||a_j|+\sum_{j=1,j\neq i}^{n}|B_{ij}||z_j|\right). \tag{3.51}$$

Proof. Given ρ in (3.47), we could integrate (3.49) and have

$$D_{ii}\rho=-\frac{sgn(z_i)}{\gamma_i}\left(\sum_{j=1}^{n}A_{ij}a_j+\sum_{j=1,j\neq i}^{n}B_{ij}z_j\right)\rho-B_{ii}\left(\frac{|z_i|}{\gamma_i}+\frac{1}{\gamma_i^2}\right)\rho+D_{ii}(0)\rho(z_i=0). \tag{3.52}$$

Here, $D_{ii}(0)\rho(z_i=0)$ could be any expression that doesn't contain z_i. We pick $D_{ii}(0)$ deliberately to guarantee that D_{ii} is non-negative. Since $B_{ii}\leq 0$ from assumption, it's sufficient to control the first term in (3.52). By letting $D_{ii}(0)=\mathcal{D}_i:=\frac{1}{\gamma_i}\left(\sum_{j=1}^{n}|A_{ij}||a_j|+\sum_{j=1,j\neq i}^{n}|B_{ij}||z_j|\right)$, we have the theorem proved. ∎

We solved the SDEs (3.43) and (3.46) numerically and computed the statistics of steady solutions as a verification of stochastic models. It is worthwhile to point out that the SDEs (3.43) and (3.46) contain an unbounded diffusion coefficient σ, which introduces a stiff problem for the numerical computations. The unboundedness reflects a mechanism for the energy to stay above a lower bound.

To resolve this numerical issue, we applied the implicit Taylor method [94], with time step size $\Delta t=5\times 10^{-4}$. Due to the uniform partition of the system, we expect the statistics to be the same for all the components of a. Here, we only exhibit the results from the locally averaged energy of the first bulk.

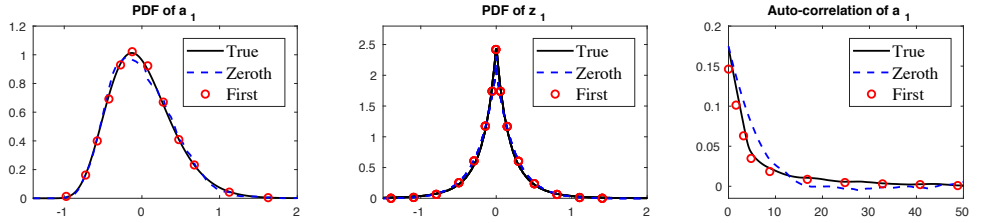

FIG. 3.6. *The figures show the approximate results of multiplicative model* (3.43) *and* (3.46)*, in terms of the steady PDF of* a_1*, steady PDF of* z_1 *and time auto correlation of* a_1*.*

The numerical results are displayed in Figure 3.6. One can see that the statistics of both a and z are recovered by both models (3.43) and (3.46). In addition, compared with models driven by additive Gaussian white noise, the multiplicative noise model has improved accuracy for the auto-correlation of a as well.

REMARK 3.3. The authors are not aware of any ergodicity result for these stochastic dynamics, especially when the coefficients are singular. Therefore, the numerical results here serve as a verification that the system will converge to the desired equilibrium measure.

REMARK 3.4. If it is just for the sake of obtaining the equilibrium density for a, one can follow a nonlinear Langevin dynamics with additive noise, *e.g.*,

$$\ddot{a} = -\nabla W(a) - \gamma\dot{a} + \sqrt{2\gamma}\zeta(t). \tag{3.53}$$

However, the corresponding solution $\dot{a}$ is still Gaussian upon equilibrium. More importantly, such a Langevin equation cannot be derived from either the Mori's or Zwanzig's projection. This is because the Mori's projection leads to linear drift terms. For the Zwanzig's projection, the memory term is no longer a time convolution, and it is not clear what systematic approximation would lead immediately to the linear damping term.

4. Summary and discussions

This work is concerned with a coarse-grained energy model directly obtained from the full molecular dynamics model. The goal is to find a more efficient model so that the heat conduction process can be obtained from studying the statistics of local energy (and flux if it's of our interest), without conducting non-equilibrium molecular dynamics simulations. Our focus has been placed on the equilibrium statistics of such models, which conceptually, is often a good starting point to develop a stochastic model. Motivated by the observation that the coarse-grained energy (and flux!) follows non-Gaussian statistics, we proposed to introduce multiplicative noise, within the Markovian embedding framework for the memory term, to ensure that the solution of the coarse-grained models has the correct equilibrium statistics.

The deterministic part of the coarse-grained models coincides with the models derived from extended thermodynamics models [3, 48]. We have established a direct connection between the coefficients in the generalized constitutive models and the statistics of the underlying microscopic models.

Although we only considered a one-dimensional model as the example, the framework is applicable to more general systems. In particular, none of the theorems assumed

the space dimensionality. The applications to nano-mechanical systems is currently underway.

REFERENCES

[1] A. Ziya Akcasu and E. Daniels, *Fluctuation analysis in simple fluids*, Phys. Rev. A, 2(3):962, 1970. 1

[2] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, Oxford University Press, 1989. 1

[3] F. X. Alvarez and D. Jou, *Memory and nonlocal effects in heat transport: From diffusive to ballistic regimes*, Appl. Phys. Lett., 90(8):083109, 2007. 1, 4

[4] J. Bonet Avalos and A. D. Mackie, *Dissipative particle dynamics with energy conservation*, Europhys. Lett., 40(2):141, 1997. 1

[5] M. Baaden and S. J. Marrink, *Coarse-grain modelling of protein-protein interactions*, Curr. Opin. Struct. Biol., 23(6):878–886, 2013. 1

[6] A. D. Baczewski and S. D. Bond, *Numerical integration of the extended variable generalized Langevin equation with a positive Prony representable memory kernel*, J. Chem. Phys., 139(4):044107, 2013. 2.3

[7] D. Bedeaux and P. Mazur, *Brownian motion and fluctuating hydrodynamics*, Physica, 76(2):247–258, 1974. 1

[8] S. Berber, Y.-K. Kwon, and D. Tománek, *Unusually high thermal conductivity of carbon nanotubes*, Phys. Rev. Lett., 84(20):4613, 2000. 1, 3.2

[9] C. Bernardin and S. Olla, *Fourier's law for a microscopic model of heat conduction*, J. Stat. Phys., 121(3-4):271–289, 2005. 1

[10] F. Bonetto, J. L. Lebowitz, and L. Rey-Bellet, *Fourier's law: A challenge for theorists*, Math. Phys., 128–150, 2000. 1

[11] C. Cattaneo, *A form of heat conduction equation which eliminates the paradox of instantaneous propagation*, Comptes Rendus, 247(4):431–433, 1958. 1, 3.2

[12] M. Ceriotti, G. Bussi, and M. Parrinello, *Colored-noise thermostats à la Carte*, J. Chem. Theo. Comput., 6(4):1170–1180, 2010. 1, 2.3

[13] C. W. Chang, D. Okawa, H. Garcia, A. Majumdar, and A. Zettl, *Breakdown of Fourier's Law in nanotube thermal conductors*, Phys. Rev. Lett., 101:075903, 2008. 1

[14] J. Che, T. Çağın, W. Deng, W. A. Goddard III, *Thermal conductivity of diamond and related materials from molecular dynamics simulations*, J. Chem. Phys., 22:6888, 2000. 1

[15] J. Chen, G. Zhang, and B. Li, *Molecular dynamics simulations of heat conduction in nanostructures: Effect of heat bath*, J. Phys. Soc. Jpn., 79:074604, 2010. 1

[16] K. Q. Chen, W. X. Li, W. Duan, Z. Shuai, and B. L. Gu, *Effect of defects on the thermal conductivity in a nanowire*, Phys. Rev. B, 72(4):045422, 2005. 1

[17] Y. Chen, D. Li, J. Yang, Y. Wu, J. R. Lukes, and A. Majumdar, *Molecular dynamics study of the lattice thermal conductivity of Kr/Ar superlattice nanowires*, Phys. B, 349(1):270–280, 2004. 1

[18] A. J. Chorin and O. H. Hald, *Stochastic Tools in Mathematics and Science*, Springer, 3, 2009. 2.1

[19] A. J. Chorin, O. H. Hald, and R. Kupferman, *Optimal prediction with memory*, Phys. D, 166(3-4):239–257, 2002. 2.1

[20] A. J. Chorin, A. P. Kast, and R. Kupferman, *Optimal prediction of underresolved dynamics*, Proc. Nat. Acad. Sci. USA, 95(8):4094–4098, 1998. 2.1, 2.1

[21] A. J. Chorin and P. Stinis, *Problem reduction, renormalization, and memory*, Comm. Appl. Math. Comput. Sci., 1:1–27, 2007. 1, 2.1

[22] B. Cockburn, S. Hou, and C.-W. Shu, *The Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. IV: The multidimensional case*, Math. Comput., 54(190):545–581, 1990. 1

[23] C. Dames and G. Chen, *Theoretical phonon thermal conductivity of Si/Ge superlattice nanowires*, J. Appl. Phys., 95(2):682–693, 2004. 1

[24] E. Darve, J. Solomon, and A. Kia, *Computing generalized Langevin equations and generalized Fokker-Planck equations*, Proc. Nat. Acad. Sci. USA, 106(27):10884–10889, 2009. 2.3

[25] P. J. Davis, *Circulant Matrices*, Amer. Math. Soc., 2013. 3.2, 3.2

[26] D. Donadio, and G. Galli, *Atomistic simulations of heat transport in silicon nanowires*, Phys. Rev. Lett., 102:195901, 2009. 1

[27] A. Donev, E. Vanden-Eijnden, A. Garcia, and J. Bell, *On the accuracy of finite-volume schemes for fluctuating hydrodynamics*, Comm. Appl. Math. Comput. Sci., 5(2):149–197, 2010. 1

[28] P. Español, *Dissipative particle dynamics with energy conservation*, Europhys. Lett., 40(6):631, 1997. 1

[29] P. Español, M. Serrano, I. Pagonabarraga, and I. Zúñiga, *Energy-conserving coarse-graining of complex molecules*, Soft Matter, 12(21):4821–4837, 2016. 1

[30] P. Español and P. Warren, *Statistical mechanics of dissipative particle dynamics*, Europhys. Lett., 30(4):191, 1995. 1

[31] W. J. Evans, L. Hu, and P. Keblinski, *Thermal conductivity of graphene ribbons from equilibrium molecular dynamics: Effect of ribbon width, edge roughness, and hydrogen termination*, Appl. Phys. Lett., 96(20):203112–203112, 2010. 1

[32] G. Faure, R. Delgado-Buscalioni, and P. Español, *The entropy of a complex molecule*, J. Chem. Phys., 146(22):224106, 2017. 3.3.2

[33] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, Academic Press, 2nd Edition, 2002. 1

[34] J. Fricks, L. Yao, T. C. Elston, and M. Gregory Forest, *Time-domain methods for diffusive transport in soft matter*, SIAM J. Appl. Math., 69(5):1277–1308, 2009. 2.3

[35] M. Fujii, X. Zhang, H. Xie, H. Ago, K. Takahashi, T. Ikuta, H. Abe, and T. Shimizu, *Measuring the thermal conductivity of a single carbon nanotube*, Phys. Rev. Lett., 95:065502, 2005. 1

[36] P. L. Garrido, P. I. Hurtado, and B. Nadrowski, *Simple one-dimensional model of heat conduction which obeys Fourier's law*, Phys. Rev. Lett., 86(24):5486, 2001. 1

[37] S. P .A. Gill, Z. Jia, B. Leimkuhler, and A. C. F. Cocks, *Rapid thermal equilibration in coarse-grained molecular dynamics*, Phys. Rev. B, 73(18):184304, 2006. 1

[38] H. Gohlke and M. F. Thorpe, *A natural coarse graining for simulating large biomolecular motion*, Biophys. J., 91(6):2115–2120, 2006. 1

[39] P. A. Golubkov and P. Ren, *Generalized coarse-grained model based on point multipole and Gay-Berne potentials*, J. Chem. Phys., 125(6):064103, 2006. 1

[40] A. Gramada and P. E. Bourne, *Coarse-graining the electrostatic potential via distributed multipoleexpansions*, Comput. Phys. Comm., 182(7):1455–1462, 2011. 1

[41] R. A. Guyer and J. A. Krumhansl, *Solution of the linearized phonon Boltzmann equation*, Phys. Rev., 148(2):766, 1966. 1

[42] A. S. Henry and G. Chen, *Spectral phonon transport properties of silicon based on molecular dynamics simulations and lattice dynamics*, J. Comput. Theor. Nanosci., 5:1–12, 2008. 1

[43] C. Hijón, M. Serrano, and P. Español, *Markovian approximation in a coarse-grained description of atomic systems*, J. Chem. Phys., 125:204101, 2006. 2.2

[44] J. Hone, M. Whitney, C. Piskoti, and A. Zettl, *Thermal conductivity of single-walled carbon nanotubes*, Phys. Rev. B, 59(4):R2514, 1999. 1

[45] J. Hu, X. Ruan, and Y. P. Chen, *Thermal conductivity and thermal rectification in graphene nanoribbons: a molecular dynamics study*, Nano Lett., 9(7):2730–2735, 2009. 1

[46] S. Izvekov and G. A. Voth, *Modeling real dynamics in the coarse-grained representation of condensed phase systems*, J. Chem. Phys., 125:151101–151104, 2006. 1

[47] K. Jolley and S. Gill, *Modelling transient heat conduction in solids at multiple length and time scales: A coupled non-equilibrium molecular dynamics/continuum approach*, J. Comput. Phys., 228(19):7412–7425, 2009. 1

[48] D. Jou, J. Casas-Vázquez, and G. Lebon, *Extended Irreversible Thermodynamics*, Springer, 1996. 1, 4

[49] M. A. Katsoulakis, A. J. Majda, and D. G. Vlachos, *Coarse-grained stochastic processes for lattice systems*, Proc. Nat. Acad. Sci. USA, 100:782–787, 2003. 1

[50] D. Kauzlarić, P. Español, A. Greiner, and S. Succi, *Markovian dissipative coarse grained molecular dynamics for a simple 2D graphene model*, J. Chem. Phys., 137(23):234103, 2012. 2.2

[51] P. Kim, L. Shi, A. Majumdar, and P. L. McEuen, *Thermal transport measurements of individual multiwalled nanotubes*, Phys. Rev. Lett., 87:215502, 2001. 1

[52] B. O. Koopman, *Hamiltonian systems and transformation in Hilbert space*, Proc. Nat. Acad. Sci. USA, 17(5):315–318, 1931. 2.1

[53] S. Kotz, T. Kozubowski, and K. Podgorski, *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance*, Springer Science and Business Media, 2012. 3.3

[54] R. Kubo, *The fluctuation-dissipation theorem*, Rep. Prog. Phys., 29(1):255–284, 1966. 2.1

[55] A. J. C. Ladd, B. Moran, and W. G. Hoover, *Lattice thermal conductivity: A comparison of molecular dynamics and anharmonic lattice dynamics*, Phys. Rev. B, 34(8):5058, 1986. 3.1

[56] H. Lei, N. A. Baker, and X. Li, *Data-driven parameterization of the generalized Langevin equation*, Proc. Nat. Acad. Sci. USA, 113(50):14183–14188, 2016. 1, 2.2, 2.2, 2.2

[57] S. Lepri, R. Livi, and A. Politi, *Heat conduction in chains of nonlinear oscillators*, Phys. Rev. Lett., 78(10):1896, 1997. 1

[58] S. Lepri, R. Livi, and A. Politi, *On the anomalous thermal conductivity of one-dimensional lattices*, Europhys. Lett., 43(3):271, 1998. 1, 3.2
[59] S. Lepri, R. Livi, and A. Politi, *Thermal conduction in classical low-dimensional lattices*, Phys. Rep., 377(1):1–80, 2003. 1, 3.2
[60] D. Li, Y. Wu, P. Kim, L. Shi, P. Yang, and A. Majumdar, *Thermal conductivity of individual silicon nanowires*, Appl. Phys. Lett., 83(14):2934–2936, 2003. 1
[61] X. Li, *A coarse-grained molecular dynamics model for crystalline solids*, Int. J. Numer. Methods Eng., 83:986–997, 2010. 1
[62] X. Li, *Heat conduction in nanoscale materials: A statistical-mechanics derivation of the local heat flux*, Phys. Rev. E, 90(3):032112, 2014. 3.1, 3.2
[63] 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(24):243128, 2015. 2.3
[64] Z. Li, H. S. Lee, E. Darve, and G. E. Karniadakis, *Computing the non-Markovian coarse-grained interactions derived from the Mori-Zwanzig formalism in molecular systems: Application to polymer melts*, J. Chem. Phys., 146(1):014104, 2017. 2.3
[65] X. Lü, W. Z. Shen, and J. H. Chu, *Size effect on the thermal conductivity of nanowires*, J. Appl. Phys., 91(3):1542–1552, 2002. 1
[66] L. Ma, X. Li, and C. Liu, *The derivation and approximation of coarse-grained dynamics from Langevin dynamics*, J. Chem. Phys., 145(20):204117, 2016. 1, 2.2, 3.3.1
[67] S. Maruyama, *A molecular dynamics simulation of heat conduction in finite length SWNTs*, Phys. B, 323(1):193–195, 2002. 1
[68] 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 Transf., 47:1783–1798, 2003. 1
[69] A. J. H. McGaughey, and M. Kaviany, *Phonon transport in molecular dynamics simulations: Formulation and thermal conductivity prediction*, Adv. Heat Transf., 39:169, 2006. 1
[70] H. Mori, *A continued-fraction representation of the time-correlation functions*, Progr. Theor. Phys., 34:399–416, 1965. 2.1, 2.1, 2.1
[71] H. Mori, *Transport, collective motion, and Brownian motion*, Progr. Theor. Phys., 33:423–450, 1965. 1, 2.1, 2.1, 2.1, 2.1
[72] S. O. Nielsen, R. E. Bulo, P. B. Moore, and B. Ensing, *Recent progress in adaptive multiscale molecular dynamics simulations of soft matter*, Phys. Chem. Chem. Phys., 12(39):12401, 2010. 1
[73] W. G. Noid, *Perspective: Coarse-grained models for biomolecular systems*, J. Chem. Phys., 139(9):090901, 2013. 1
[74] W. G. Noid, J.-W. Chu, G. S. Ayton, V. Krishna, S. Izvekov, G. A. Voth, A. Das, and H. C. Andersen, *The multiscale coarse-graining method. I. A rigorous bridge between atomistic and coarse-grained models*, J. Chem. Phys., 128(24):244114, 2008. 1
[75] S. Nosé, *A molecular dynamics method for simulations in the canonical ensemble*, Mol. Phys., 52(2):255–268, 1984. 1
[76] M. A. Osman and D. Srivastava, *Temperature dependence of the thermal conductivity of single-wall carbon nanotubes*, Nanotechnology, 12(1):21, 2001. 1
[77] G. A. Pavliotis, *Stochastic Processes and Applications: Diffusion Processes, the Fokker-Planck and Langevin Equations*, Springer, 60, 2014. 2.2, 3.3.1
[78] E. Pop, D. Mann, Q. Wang, K. Goodson, and H. Dai, *Thermal conductance of an individual single-wall carbon nanotube above room temperature*, Nano Lett., 6:96–100, 2006. 1
[79] P. Poulain, A. Saladin, B. Hartmann, and C. Pravost, *Insights on protein-DNA recognition by coarse grain modelling*, J. Comput. Chem., 29(15):2582–2592, 2008. 1
[80] M. Praprotnik, L. Delle Site, and K. Kremer, *Multiscale simulation of soft matter: From scale bridging to adaptive resolution*, Annu. Rev. Phys. Chem., 59(1):545–571, 2008. 1
[81] S. Riniker, J. R. Allison, and W. F. van Gunsteren, *On developing coarse-grained models for biomolecular simulation: a review*, Phys. Chem. Chem. Phys., 14(36):12423, 2012. 1
[82] M. Ripoll, P. Español, and M. H. Ernst, *Dissipative particle dynamics with energy conservation: Heat conduction*, Int. J. Mod. Phys. C, 9(08):1329–1338, 1998. 1
[83] H. Risken and H. Haken, *The Fokker-Planck Equation: Methods of Solution and Applications*, Second Edition, Springer, 1989. 3.3.1, 3.3.1
[84] R. E. Rudd and J. Q. Broughton, *Coarse-grained molecular dynamics and the atomic limit of finite element*, Phys. Rev. B, 58(10):5893–5896, 1998. 1
[85] R. E. Rudd and J. Q. Broughton, *Coarse-grained molecular dynamics: Nonlinear finite elements and finite temperature*, Phys. Rev. B, 72:144104, 2005. 1
[86] J. F. Rudzinski, and W. G. Noid, *The role of many-body correlations in determining potentials for coarse-grained models of equilibrium structure*, J. Phys. Chem. B, 116(29):8621–8635,

2012. 1

[87] A. V. Savin, Y. S. Kivshar, and B. Hu, *Suppression of thermal conductivity in graphene nanoribbons with rough edges*, Phys. Rev. B, 82(19):1954, 2010. 1

[88] P. K. Schelling, S. R. Phillpot, and P. Keblinski, *Comparison of atomic-level simulation methods for computing thermal conductivity*, Phys. Rev. B, 65(14):144306, 2002. 1, 3.2

[89] Q. Shi, P. Liu, and G. A. Voth, *Coarse-graining in interaction space: An analytical approximation for the effective short-ranged electrostatics*, J. Phys. Chem. B, 112(50):16230–16237,2008. 1

[90] S. Silling and R. B. Lehoucq, *Statistical coarse-graining of molecular dynamics into peridynamics*, Technical Report, Sandia National Laboratories, SAND2007-6410, 2007. 1

[91] A. Singh and E. B. Tadmor, *Thermal parameter identification for non-Fourier heat transfer from molecular dynamics*, J. Comput. Phys., 299:667–686, 2015. 1

[92] H. Spohn, *Nonlinear fluctuating hydrodynamics for anharmonic chains*, J. Comput. Chem., 154(5):1191–1227, 2014. 3.2

[93] M. Stepanova, *Dynamics of essential collective motions in proteins: Theory*, Phys. Rev. E, 76:051918, 2007. 1

[94] T. Tian and K. Burrage, *Implicit Taylor methods for stiff stochastic differential equations*, Appl. Numer. Math., 38(1-2):167–185, 2001. 3.3.2

[95] M. Toda, R. Kubo, and N. Hashitsume, *Statistical Physics II. Nonquilibrium Statistical Mechanics*, Springer, 1983. 2.1

[96] D. Y. Tzou, *Experimental support for the lagging behavior in heat propagation*, J. Thermophys. Heat Transf., 9(4):686–693,1995. 1, 3.2

[97] D. Y. Tzou, *Nonlocal behavior in phonon transport*, Int. J. Heat Mass Transf., 15:475–481, 2011. 1

[98] P. Vernotte, *Paradoxes in the continuous theory of the heat equation*, CR Acad. Sci., 246(3):3154–3155, 1958. 1, 3.2

[99] A. N. Volkov and L. V. Zhigilei, *Heat conduction in carbon nanotube materials: Strong effect of intrinsic thermal conductivity of carbon nanotubes*, Appl. Phys. Lett., 101:043113, 2012. 1

[100] S. G. Volz and G. Chen, *Molecular dynamics simulation of thermal conductivity of silicon nanowires*, Appl. Phys. Lett., 75(14):2056–2058, 1999. 1, 3.1

[101] S. G. Volz and G. Chen, *Molecular-dynamics simulation of thermal conductivity of silicon crystals*, Phys. Rev. B, 61:2651, 2000. 1

[102] J.-S. Wang, *Quantum thermal transport from classical molecular dynamics*, Phys. Rev. Lett., 99:160601, 2007. 1

[103] L. Wang, B. Hu, and B. Li, *Logarithmic divergent thermal conductivity in two-dimensional nonlinear lattices*, Phys. Rev. E, 86:040101, 2012. 1

[104] S.-C. Wang, X.-G. Liang, X.-H. Xu, and T. Ohara, *Thermal conductivity of silicon nanowire by nonequilibrium molecular dynamics simulations*, J. Appl. Phys., 105(1):014316–014316, 2009. 1

[105] Z. J. Yang, X. Wu, and X. Li, *A generalized Irving–Kirkwood formula for the calculation of stress in molecular dynamics models*, J. Chem. Phys., 137(13):134104, 2012. 3.1

[106] H. Yoshida, *Construction of higher order symplectic integrators*, Phys. Lett. A, 150(5-7):262–268, 1990. 3.2

[107] C. Yu, L. Shi, Z. Yao, D. Li, and A. Majumdar, *Thermal conductance and thermopower of an individual single-wall carbon nanotube*, Nano Lett., 5(9):1842–1846, 2005. 1

[108] Z. Zhang, L. Lu, W. G. Noid, V. Krishna, J. Pfaendtner, and G. A. Voth, *A systematic methodology for defining coarse-grained sites in large biomolecules*, Biophys. J., 95(11):5073–5083, 2008. 1

[109] R. Zwanzig, *Nonlinear generalized Langevin equations*, J. Comput. Chem., 9:215–220, 1973. 1, 2.1

[110] R. Zwanzig, *Nonequilibrium Statistical Mechanics*, Oxford University Press, 2001. 2.1, 2.1