Physics Area
Ph.D. course in Theory and Numerical Simulation of Condensed Matter

THERMOELECTRIC EFFECTS IN POLAR LIQUIDS AND IONIC CONDUCTORS

Candidate:
Enrico Drigo

Supervisor:
Prof. Stefano Baroni

Academic year 2023/2024

A Pia, Paolo e Mariarosa

ABSTRACT

Transport theory describes non-equilibrium conditions. The Onsager equations set a linear relationship between thermodynamic forces and macroscopic currents. When a temperature gradient is applied, it can generate heat and ions diffusion. Similarly, an electric field can induce electric currents and heat transport. The off-diagonal elements of the Onsager matrix address coupled, e. g. thermoelectric, effects. In the 1950s, Green and Kubo developed a sound equilibrium framework for transport phenomena based on linear response theory. Despite the theoretical depth, computational inefficiencies hindered its broad application to equilibrium molecular dynamics simulations. In the last decade, the introduction of gauge and convective invariance principles addressed common misconceptions in the Green Kubo theory, concerning thermoelectric and heat transport. The aim of this Thesis is threefold.

- I investigate thermoelectric effects in classical fluids. In insulating liquids, thermopolarization is a quasi-equilibrium phenomenon. After being perturbed by a temperature inhomogeneity, the dielectric degrees of freedom relax on a microscopic time scale. In conducting fluids, a temperature gradient induces electric currents. The Seebeck effect refers to the electric field that must be applied to counterbalance the temperature gradient and then achieve zero net charge flow. We study thermoelectric transport in several ionic conductors.

- We develop a Bayesian strategy to compute both diagonal and off-diagonal Onsager coefficients from equilibrium molecular dynamics simulations. The Bayesian approach leverages the statistical properties of Green-Kubo estimators. A single statistical model can be designed to estimate the entire Onsager matrix.

- Finally, I participated in a collaboration studying heat transport in glasses. In disordered solids, the acoustic contribution to thermal conductivity is challenging to estimate. The investigation of nearly 140000-atoms harmonic models of glasses, reveals the presence of Rayleigh scattering in the low-frequency regime. In harmonic disordered systems, our findings predict a divergent bulk thermal conductivity at all temperatures. The divergence is cured by proper accounting for anharmonic effects.

The results presented in this Thesis are discussed in five papers, four published and one under peer review:

1. Enrico Drigo and Stefano Baroni. “Seebeck Coefficient of Liquid Water from Equilibrium Molecular Dynamics.” In: *Journal of Chemical Theory and Computation* 19.23 (2023). PMID: 37823575, pp. 8855–8860. DOI: 10.1021/acs.jctc.3c00760.

   I propose a quasi-equilibrium theoretical framework to study the thermopolarization effect in polar insulating liquids from the long-wavelength limit of a suitable static susceptibility. The long-wavelength limit is estimated via Bayesian linear regression analysis, which also allows to evaluate the uncertainty associated to the extrapolation. The method is tested on liquid water and benchmarked against state-of-the-art non-equilibrium simulations.

2. Enrico Drigo, Stefano Baroni, and Paolo Pegolo. “Seebeck Coefficient of Ionic Conductors from Bayesian Regression Analysis.” In: *Journal of Chemical Theory and Computation* 20.14 (2024). PMID: 38856670, pp. 6152–6159. DOI: 10.1021/acs.jctc.4c00124.

   We investigate the Seebeck effect in ionic conductors and develop a Bayesian regression analysis to address thermoelectric transport coefficients. The technique is benchmarked against standard Green-Kubo methods on molten salts in a wide range of pressure-temperature conditions.

3. Paolo Pegolo, Enrico Drigo, Federico Grasselli, and Stefano Baroni. "Transport coefficients from equilibrium molecular dynamics". Submitted to: *The Journal of Chemical Physics*. 2024.

   The work is an invited paper in *The Journal of Chemical Physics*, featured under the special topic: *Molecular Dynamics, Methods and Applications 60 Years after Rahman*. In this paper we review spectral methods for computing transport coefficients. On the theoretical side, we introduce a unified Bayesian model to address the entire Onsager matrix; on the practical side, we apply Bayesian regression analysis to estimate the transport coefficients of a promising Li-based solid-state electrolyte.

4. Enrico Drigo, Maria Grazia Izzo, and Stefano Baroni. “Heat conductivity from energy-density fluctuations.” In: *The Journal of Chemical Physics* 159.18 (2023), p. 184107. ISSN: 0021-9606. DOI: 10.1063/5.0168732.

   Building on hydrodynamic considerations, we investigate energy density fluctuations and compute thermal conductivity from their long-wavelength behavior. The energy density approach is compared to a recently proposed similar technique, which instead utilizes mass density fluctuations.

5. Alfredo Fiorentino, Enrico Drigo, Stefano Baroni, and Paolo Pegolo. “Unearthing the foundational role of anharmonicity in heat transport in glasses.” In: *Phys. Rev. B* 109 (22 2024), p. 224202. DOI: 10.1103/PhysRevB.109.224202.

In the presence of disorder-induced Rayleigh scattering, harmonic theories of heat transport in glasses inevitably predict that bulk thermal conductivity diverges. We demonstrate that, disorder alone is insufficient to yield a finite bulk thermal conductivity in glasses. Anharmonic effects are crucial to provide physically sensible results in disordered systems. Our findings are tested through extensive numerical simulations.

CONTENTS

INTRODUCTION 1

1 THEORY OF TRANSPORT PROCESSES 3
- 1.1 Irreversible thermodynamics 3
- 1.2 Green-Kubo response theory 6
  - 1.2.1 Linear response theory 6
  - 1.2.2 Representation of the heat flux 12
- 1.3 Invariance principles in the theory of transport 14
  - 1.3.1 Gauge invariance principle of transport coefficients 14
  - 1.3.2 Convective invariance of heat conductivity in multicomponent systems 15
- 1.4 Hydrodynamic description of a diffusive process 16

2 CEPSTRAL AND BAYESIAN REGRESSION ANALYSIS OF RESPONSE FUNCTIONS 19
- 2.1 Statistical properties of Green-Kubo estimators 19
- 2.2 Cepstral analysis of diagonal Onsager coefficients 22
  - 2.2.1 Cepstral analysis 23
  - 2.2.2 Numerical experiments on cepstral analysis 25
- 2.3 Bayesian regression analysis of Wishart processes 29
  - 2.3.1 Bayesian regression analysis 30
  - 2.3.2 Numerical experiments on Bayesian regression analysis 31
- 2.4 Bayesian extrapolation of static response functions 34

3 SEEBECK COEFFICIENT OF CLASSICAL FLUIDS 39
- 3.1 Thermopolarization effect in insulating polar fluids 39
  - 3.1.1 Thermopolarization effect from static response theory 39
  - 3.1.2 Bayesian regression analysis of the thermopolarization coefficient of liquid water 40
- 3.2 Thermoelectric effects in ionic conductors 45
  - 3.2.1 Seebeck effect and thermoelectric transport in ionic conductors 45
  - 3.2.2 Bayesian regression analysis of the Seebeck coefficient 46
  - 3.2.3 Bayesian regression analysis of transport coefficients 51

4 THERMAL TRANSPORT IN SOLID AND LIQUID INSULATORS 57
- 4.1 Extrapolation of thermal conductivity from energy density fluctuations 57
- 4.2 Foundations of Quasi-Harmonic-Green-Kubo theory 63

4.3 The role of anharmonicity in heat transport in glasses 66
CONCLUSIONS 76

Appendix 77
A KUBO'S AND MARTIN'S DISCUSSION ON TRANSPORT THEORY 79
B MARTIN'S ANALYSIS ON RESPONSES TO ELECTRIC PERTURBATIONS 81
C DEBENEDETTI CALCULATION OF PARTIAL MOLAR QUANTITIES 87
D VALIDATION OF THE MACHINE LEARNING INTERATOMIC POTENTIAL OF SOLID-STATE ELECTROLYTES 89
ACKNOWLEDGMENTS 92

BIBLIOGRAPHY 93

ACRONYMS

GK	Green Kubo
EMD	Equilibrium Molecular Dynamics
NEMD	Non Equilibrium Molecular Dynamics
HE	Helfand Einstein
PSD	Power Spectral Density
AIC	Akaike Information Criterion
MCMC	Monte Carlo Markov Chain
PBC	Periodic Boundary Conditions
BIC	Bayesian Information Criterion
MMSE	Minimum Mean Square Error
KK	Kramers Kronig
DF	Dielectric Function
NLL	Negative Log-Likelihood
SSE	Solid State Electrolyte
NEP	Neuroevolution Potential
DFT	Density Functional Theory
WF	Window Filter
LP	Landau Planczek
EDSF	Energy Density Structure Factor
IFC	Interatomic Force Constants
QHGK	Quasi Harmonic Green Kubo
FGR	Fermi Golden Rule
BTE	Boltzmann Transport Equation
AF	Allen Feldman
RTA	Relaxation Time Approximation
WBTE	Wigner Boltzmann Transport Equation
VDSF	Vibrational Dynamical Structure Factor
LPS	Li_3PS_4

Chi cerca, trova; chi ricerca, ritrova.

— Ennio De Giorgi —

INTRODUCTION

The modern theory of transport processes stands on the shoulders of giants: Albert Einstein, Lars Onsager, Hendrik Casimir, Herbert B. Callen, Ryogo Kubo, Melville S. Green, John G. Kirkwood, Joaquin M. Luttinger, Hazime Mori, Paul C. Martin— just to name a few.

The regression of fluctuations is encoded in the second law of thermodynamics. The theoretical analysis of relaxation effects has been enormously successful, with broad applications, ranging from theoretical physics to industry and technology. Being so flexible, transport theory has many different applications, including planetary science, design of energy storage devices and heat management. In the 1930s, Onsager established the celebrated *reciprocity relations*. The groundbreaking work expanded the results published in 1905 by Einstein on Brownian motion. In a linear response theory approach, Green and Kubo derived the time-honored Green-Kubo formulas of transport coefficients. Combined with the development of numerical simulations of realistic atomistic systems, transport theory allows to study non-equilibrium properties of materials from equilibrium molecular dynamics simulations. The theoretical benefits notwithstanding, computational limitations and conceptual misconceptions penalized equilibrium methods. In response to these drawbacks, non-equilibrium molecular dynamics approaches were introduced. In non-equilibrium simulations one superimposes thermodynamic forces on the simulation cell, directly mimicking transport phenomena. Despite the adherence to the actual process, non-equilibrium methods are sensitive to size effects. In this Thesis, I study equilibrium techniques, aiming to address some of their numerical and theoretical challenges.

In the last decade, significant advancements in the theory of transport have investigated efficient equilibrium approaches, thermoelectric effects and lattice theory of heat transport. Remarkably, the conundrum concerning the arbitrary definition of the densities was finally solved. Densities, and the associated currents, depend on the microscopic partitioning of an extensive conserved quantity into local contributions. The arbitrariness hindered the extension of the theory of transport to *ab-initio* simulations. The *gauge invariance* principle of transport coefficients demonstrates that the Onsager matrix is independent of the local definition of the densities. Recently, *cepstral analysis* partially lifted the computational burden of the computation of diagonal Onsager coefficients, such as: thermal conductivity, electric conductivity, ionic diffusivities and, shear and bulk viscosity. In speech recognition, cepstral analysis is a common tool used to filter the noise affecting the signal. Leveraging the statistical properties of diagonal Green-Kubo estimators, cepstral analysis effectively

reduces the numerical cost of accurate transport coefficients and reasonable error bars, relying on relatively short equilibrium molecular dynamics simulations. The computational benefits notwithstanding, cepstral analysis cannot be applied to off-diagonal elements of the Onsager matrix.

Coupled phenomena, e. g. thermoelectric effects, are crucial for energy storage devices and heat management. In ionic conductors and polar fluids, temperature gradients induce electric currents and electric polarization, respectively. Among thermoelectric effects, the Seebeck's is one of the most important: it describes the interplay between temperature inhomogeneities and dielectric degrees of freedom. Thanks to transport theory, in insulating fluids the Seebeck coefficient can be written as the ratio of two vanishing Onsager coefficients, making practical evaluations impossible. Tackling thermopolarization of polar fluids, and the consequent indeterminate form $\frac{0}{0}$, requires a paradigm shift. In this Thesis, developing Bayesian regression tools, thermoelectric effects in insulating fluids and ionic conductors are directly investigated. In solids, under the hypothesis that the interatomic distances are much larger than vibrations, heat carriers can be described as quasi-particles modes in the harmonic approximation, i. e. *phonons*. Progress on heat transport, thanks to the development of the Quasi-Harmonic-Green-Kubo theory and of the Wigner Boltzmann transport equation, has been fundamental to address both crystalline and disordered systems. This Thesis delves into the theory of transport in classical fluids as well as in solids. The Thesis is structured as follows.

Chapter 1 introduces the general concepts of the theory of transport phenomena. An overview of the most common approaches is presented. The recent gauge and convective invariance principles are discussed. Chapter 2 is devoted to numerical methods. I highlight the numerical drawbacks of equilibrium approaches and propose spectral methods to enhance their efficiency. Special attention is spent on Bayesian regression analysis. In Chapter 3, theoretical and numerical results on thermopolarization in liquid water are presented. I also analyze the Seebeck coefficient of ionic conductors, applying a recently proposed Bayesian scheme. The statistical properties of the Green-Kubo estimators lead to an all-encompassing Bayesian regression protocol, which evaluates the entire Onsager matrix from a single statistical model. The protocol is applied to estimate the transport coefficients of a promising Li-based solid-state electrolyte. Finally, in Chapter 4, I review some recent advances on lattice heat transport in glasses, namely the Quasi-Harmonic-Green-Kubo theory. The foundational role of anharmonic effects is discussed. Remarkably, in harmonic disordered systems, the inclusion of only Rayleigh scattering results in a divergent bulk thermal conductivity at all temperatures.

1

THEORY OF TRANSPORT PROCESSES

TRANSPORT THEORY is of particular importance both for its theoretical depth and for its practical applications. Quasi-equilibrium phenomena are investigated under the hypotheses of microscopic reversibility and linearity of regressions. This chapter presents the Green's and Kubo's linear response theory approach, in its classical and quantum formulation, as developed in a series of seminal papers in the 1950s. The dynamical response to time-space dependent perturbations is compared with the phenomenological Fick's law. Additionally, I review some invariance principles of transport coefficients and the hydrodynamic behavior of conserved densities in normal fluids.

1.1 IRREVERSIBLE THERMODYNAMICS

An equilibrium state is characterized by a set of extensive thermodynamic variables, such as energy, volume and the number of particles, $\{E; V; N\}$, which determine entropy $\mathcal{S}(E, V, N)$. The derivatives of entropy are intensive properties of the state, such as temperature, pressure and chemical potential, $\left\{\frac{1}{T}; \frac{P}{T}; -\frac{\mu}{T}\right\}$ [1]. In a bipartite system, I define the subsystems' entropy and their respective extensive variables as $\mathcal{S}(\{A_k\})$, $\mathcal{S}'\left(\{A'_k\}\right)$ [1]. At equilibrium, the total entropy $\mathcal{S}^\circ$ is maximized under the constrain $A^\circ_k = A_k + A'_k$ [1],

$$\mathcal{F}_k \equiv \left.\frac{\partial \mathcal{S}^\circ}{\partial A_k}\right|_{A^\circ_k} = 0, \tag{1.1}$$

where $\mathcal{F}_k$ are the so-called *affinities*. Since $\mathcal{S}^\circ$ is extensive, it is the sum of the subsystems' contributions [1]:

$$\frac{\partial \mathcal{S}}{\partial A_k} \equiv f_k \tag{1.2}$$

$$\left.\frac{\partial \mathcal{S}^\circ}{\partial A_k}\right|_{A^\circ_k} = \left.\frac{\partial\left(\mathcal{S} + \mathcal{S}'\right)}{\partial A_k}\right|_{A^\circ_k} = \frac{\partial \mathcal{S}}{\partial A_k} - \frac{\partial \mathcal{S}'}{\partial A'_k}. \tag{1.3}$$

Combining Eq. (1.1) and (1.2), the affinity reads [1]

$$\mathcal{F}_k = f_k - f'_k. \tag{1.4}$$

In thermodynamic equilibrium, f_ks are constant across the subsystems. In *quasi-equilibrium* conditions, $\mathcal{F}_k$ is non zero and entropy evolves in time [1]. The quasi-equilibrium condition can be rationalized as the perturbation of a system at equilibrium by a fluctuation, $\boldsymbol{\alpha} \equiv \{\alpha_k\}$, of some extensive quantity

A_k, such that: $A_k = A_k^\circ + \alpha_k$ [2, 3]. Entropy is a function of the fluctuations, $\mathcal{S}(\boldsymbol{\alpha})$, and is maximized and redefined to vanish at equilibrium: $\mathcal{S}^\circ \equiv \mathcal{S}(\boldsymbol{\alpha} = 0)$ [2, 3]. The affinities, $\mathcal{F}_k \equiv \frac{\partial \mathcal{S}(\alpha)}{\partial \alpha_k}$, measure the tendency of the system to restore equilibrium and characterize *irreversible* processes [2, 3]:

$$\begin{aligned} \mathcal{S} &= \sum_k \alpha_k \mathcal{F}_k \\ \dot{\mathcal{S}}(t) &= \sum_k \dot{\alpha}_k(t) \mathcal{F}_k. \end{aligned} \tag{1.5}$$

From the regression of $\boldsymbol{\alpha}(t)$ towards equilibrium, one can deduce the evolution of the rate of production of entropy. At the lowest order in the fluctuations, the variation of $\mathcal{S}$ can be expanded as [2, 3]

$$\mathcal{S} = -\frac{1}{2} \sum_{ij} S_{ij} \alpha_i \alpha_j. \tag{1.6}$$

In quasi-equilibrium processes, a linear relationship is assumed to hold between the fluctuations of the extensive variables and the affinities [1, 3–5],

$$\dot{\alpha}_i = \sum_k L_{ik} \mathcal{F}_k \tag{1.7}$$

where $\mathcal{F}_i = \sum_k R_{ik} \dot{\alpha}_k$ and the matrices $\mathbf{L}$, $\mathbf{R}$ are mutually reciprocal [2]. Thanks to Eq. (1.7) and Eq. (1.6), the affinities are:

$$\begin{aligned} \mathcal{F}_i &= \sum_j S_{ij} \alpha_j \\ \left\langle \alpha_i \mathcal{F}_j \right\rangle_0 &= \delta_{ij}, \end{aligned} \tag{1.8}$$

and $\langle \circ \rangle_0$ is the expectation value computed over the microcanonical ensemble [3]. For simplicity, I discuss fluctuations of variables that are even functions of the velocities and that are thus time-reversal invariant. The future behavior of the system, having fixed all fluctuations $\boldsymbol{\alpha}(t)$ at instant t, is identical to its past behavior [3]:

$$\left\langle \alpha_i(t+\tau) \right\rangle_{\boldsymbol{\alpha}(t)} = \left\langle \alpha_i(t-\tau) \right\rangle_{\boldsymbol{\alpha}(t)}, \tag{1.9}$$

where the average, $\langle \circ \rangle_{\boldsymbol{\alpha}(t)}$, runs over the portion of the microcanonical ensemble on which $\boldsymbol{\alpha}(t)$ is constant [3]. Multiplying by $\alpha_k(t)$ and computing the expectation value over the total microcanonical ensemble, I obtain the *microscopic reversibility condition*, that is fundamental to the forthcoming discussion on transport phenomena [1, 3–5],

$$\left\langle \alpha_i(t+\tau) \alpha_k(t) \right\rangle_0 = \left\langle \alpha_i(t-\tau) \alpha_k(t) \right\rangle_0. \tag{1.10}$$

The final assumption refers to the dynamics of the fluctuations. There exists a time-scale, τ, over which the fluctuations evolve as [3]

$$\begin{aligned} \left\langle \alpha_i(t+\tau) - \alpha_i(t) \right\rangle_{\boldsymbol{\alpha}(t)} &= \tau \sum_k L_{ik} \mathcal{F}_k, \\ \left\langle \alpha_k(t) \left(\alpha_i(t+\tau) - \alpha_i(t) \right) \right\rangle_0 &= \tau L_{ik}. \end{aligned} \tag{1.11}$$

The time-scale τ is sufficient for the system to reach a steady flow, yet is much shorter than the equilibration time after which the disturbance is appreciably reduced [3]. Because of microscopic reversibility and time-reversal symmetry, the matrix $\mathbf{L}$ reads [3]

$$\begin{aligned}\langle \alpha_k(t)\,(\alpha_i(t+\tau)-\alpha_i(t))\rangle_0 &= \tau L_{ik}\\ \langle \alpha_i(t)\,(\alpha_k(t+\tau)-\alpha_k(t))\rangle_0 &= \tau L_{ki}.\end{aligned} \tag{1.12}$$

Eq. (1.11) and Eq. (1.7) are related to the popular Onsager phenomenological equations which are generally written in terms of the currents, $\boldsymbol{J}_i$, induced by the *thermodynamic forces*, i. e. the gradients of the intensive variables conjugated to the extensive ones being conserved, ∇f_k:

$$\boldsymbol{J}_i = \sum_k L_{ik}\nabla f_k. \tag{1.13}$$

Under the assumption of microscopic reversibility and linearity of the regressions, for the matrix $\mathbf{L}$, which enters in Eqs. (1.11) (1.7) (1.13), the *Onsager reciprocity relations* hold [3–5]

$$\mathbf{L} = \mathbf{L}^T. \tag{1.14}$$

Here, I present a direct derivation of Eq. (1.13), connecting fluxes and thermodynamic forces [1]. The differential forms of entropy and its per unit volume density, s, read:

$$\begin{aligned}dS &= \sum_k f_k\, dA_k\\ ds &= \sum_k f_k\, da_k\,.\end{aligned} \tag{1.15}$$

In continuous systems, where each subsystem can be labeled by a continuous variable $\mathbf{r}$ —dropped for clarity— s does not obey the continuity equation of conserved densities [1]:

$$\begin{aligned}\dot{s} &= \frac{\partial s}{\partial t} + \nabla\cdot\boldsymbol{J}_S\\ 0 &= \frac{\partial a_k}{\partial t} + \nabla\cdot\boldsymbol{J}_k.\end{aligned} \tag{1.16}$$

The flux of a conserved quantity, $\boldsymbol{J}_k$, is $\frac{dA_k}{dt}$ [1]. Instead, the flux of entropy is defined from Eq. (1.15) as $\boldsymbol{J}_S = \sum_k f_k \boldsymbol{J}_k$ [1]. According to the definition of the entropy flux, the entropy production rate is [1]:

$$\begin{aligned}\dot{s} &= \frac{\partial s}{\partial t} + \sum_k \nabla\cdot f_k\boldsymbol{J}_k\\ \frac{\partial s}{\partial t} &= \sum_k f_k\frac{\partial a_k}{\partial t} = -\sum_k f_k\nabla\cdot\boldsymbol{J}_k,\end{aligned} \tag{1.17}$$

and therefore I obtain the usual relation

$$\dot{s} = \sum_k \boldsymbol{J}_k \cdot \nabla f_k. \tag{1.18}$$

Eqs. (1.18) and Eq. (1.5) link the rate of increase of entropy to the affinities and the fluctuations. The affinities, that are defined as the difference of the intensive thermodynamic variables in discrete systems, are the gradients of the same intensive variables in continuous systems [1]. This formulation is more common than the one obtained from the regression of fluctuations and, under the hypothesis of linearity between fluxes and forces, leads directly to the *Onsager phenomenological equations* [1], Eq. (1.13),

$$\boldsymbol{J}_i = \sum_k L_{ik} \nabla f_k.$$

1.2 GREEN-KUBO RESPONSE THEORY

1.2.1 *Linear response theory*

In a series of seminal papers [6–9], Green and Kubo derived a sound theoretical framework for the description of transport processes from the time-fluctuations of dynamical variables associated to irreversible processes. The goal is to construct a statistical-mechanical theory for transport phenomena near thermal equilibrium. For example, thanks to the fluctuation dissipation theorem, one can compute the friction constant of a Brownian particle from the correlation of the forces acting on it [10, 11]. The methodology is restricted to perturbations that can be expressed as an additional term in the Hamiltonian. Even if temperature gradients can not be directly included in the Hamiltonian, the theory also applies to heat transport. In addressing heat transport, I rely on so-called *mechanical proxies* for thermal perturbations, as explained by Luttinger in 1964 [12].

Following the discussion presented in Ref. [6], let us consider an isolated system, described by a Hamiltonian $\mathcal{H}$, and an external time dependent force, $F(t)$, acting on it that is coupled to a quantity A. The force induces a perturbation in the Hamiltonian [6, 7]:

$$\mathcal{H}'(\Gamma, t) = -A(\Gamma)F(t). \tag{1.19}$$

where Γ is the phase-space variable. The change of a quantity $\Delta B(t)$, induced by $\mathcal{H}'$ is determined by the equilibrium properties of $\mathcal{H}$. The statistical ensemble is defined by the distribution function, f, which depends on the phase-space variable, Γ. The ensemble average, $\langle \circ \rangle_0$, of an observable A is:

$$\langle A \rangle_0 = \int f(\Gamma) A(\Gamma)\, d\Gamma\,. \tag{1.20}$$

A generic observable, A, evolves according to the Hamiltonian [6, 7]

$$\dot{\hat{A}} = \left(\hat{\mathcal{H}}, \hat{A}\right), \tag{1.21}$$

where I indicate $\widehat{A} \equiv A(\Gamma)$ and $\left(\widehat{A}, \widehat{B}\right) = \sum_i \frac{\partial \widehat{A}}{\partial \mathbf{R}_i} \frac{\partial \widehat{B}}{\partial \mathbf{P}_i} - \frac{\partial \widehat{A}}{\partial \mathbf{P}_i} \frac{\partial \widehat{B}}{\partial \mathbf{R}_i}$ is the usual Poisson bracket and $\mathbf{R}_i$, $\mathbf{P}_i$ are the canonical position and momentum of atom i, respectively. At equilibrium, one has $\frac{\partial \widehat{f}}{\partial t} = \left(\widehat{\mathcal{H}}, \widehat{f}\right) = 0$. The force $F(t)$ is switched on in the infinite past, $t = -\infty$, and adiabatically perturbs the distribution function [6, 7]. At linear order in the perturbed distribution, $\widehat{f}(t) = \widehat{f} + \widehat{f}'(t)$, one has

$$\begin{aligned} \frac{\partial \widehat{f}(t)}{\partial t} &= \left(\widehat{\mathcal{H}}, \widehat{f}(t)\right) + \left(\widehat{\mathcal{H}}', \widehat{f}\right) \\ \frac{\partial \widehat{f}'(t)}{\partial t} &= \left(\widehat{\mathcal{H}}, \widehat{f}'(t)\right) - F(t)\left(\widehat{A}, \widehat{f}\right), \end{aligned} \tag{1.22}$$

provided that $f(-\infty) = f$. The variation of the equilibrium distribution is integrated in time as [6, 7],

$$\widehat{f}'(t) = -\int_{-\infty}^{t} \exp\{i(t-t')L\}\left(\widehat{A}, \widehat{f}\right) F(t')\, dt', \tag{1.23}$$

where $\left(\widehat{\mathcal{H}}, \circ\right) \equiv -iL$ and L is the Liouvillian time evolution operator such that $\dot{\widehat{B}} = \left(\widehat{B}, \widehat{\mathcal{H}}\right)$. One can verify that Eq. (1.23) is a solution of Eq. (1.22), by simple substitution and applying the chain rule. The change of B, $\langle \Delta B(t) \rangle$, is [6, 7]

$$\begin{aligned} \left\langle \Delta \widehat{B}(t) \right\rangle &\equiv \int \widehat{f}'(t) \widehat{B}\, d\Gamma \\ &= -\int d\Gamma \int_{-\infty}^{t} \exp\{i(t-t')L\}\left(\widehat{A}, \widehat{f}\right) \widehat{B} F(t')\, dt' \\ &= -\int d\Gamma \int_{-\infty}^{t} \left(\widehat{A}, \widehat{f}\right) \widehat{B}(t-t') F(t')\, dt'. \end{aligned} \tag{1.24}$$

The *response function*, $\phi_{BA}(t)$, is defined as [6, 7]

$$\phi_{BA}(t) = -\int \left(\widehat{A}, \widehat{f}\right) \widehat{B}(t)\, d\Gamma, \tag{1.25}$$

and describes completely the change of the quantity B due to the external force F coupled to an observable A,

$$\left\langle \Delta \widehat{B}(t) \right\rangle = \int_{-\infty}^{t} \phi_{BA}(t-t') F(t')\, dt'. \tag{1.26}$$

Under the assumption that the ensemble is *canonical* and thanks to the cyclic properties of the Poisson brackets, one has [6, 7]

$$\begin{aligned} \phi_{BA}(t) &= -\int \left(\widehat{A}, \widehat{f}\right) \widehat{B}(t)\, d\Gamma \\ &= \beta \int \widehat{f}\left(\widehat{A}, \widehat{\mathcal{H}}\right) \widehat{B}(t)\, d\Gamma = \beta \int \widehat{f} \dot{\widehat{A}} \widehat{B}(t)\, d\Gamma \\ &= \beta \int d\Gamma\, \widehat{f} \int_{-\infty}^{t} \widehat{A}\left(\widehat{\mathcal{H}}, \widehat{B}\right) \exp\{i\,(t-t')\,L\}\, dt' = -\beta \int \widehat{f} \widehat{A} \dot{\widehat{B}}(t)\, d\Gamma, \end{aligned} \tag{1.27}$$

where $\beta = \frac{1}{k_B T}$, T is the temperature of the system and k_B is the Boltzmann constant. Therefore, the response function can be evaluated as an equilibrium expectation value:

$$\begin{aligned}\phi_{BA}(t) &= \beta \left\langle \dot{\widehat{A}} \widehat{B}(t) \right\rangle_0 \\ &= -\beta \left\langle \widehat{A} \dot{\widehat{B}}(t) \right\rangle_0 .\end{aligned} \tag{1.28}$$

The perturbation, $F(t) = e^{\epsilon t} F \Theta(-t)$, is adiabatically switched on at $-\infty$ and removed at $t = 0$, where $\Theta(-t)$ is the Heaviside function. In the limit of $\epsilon \to 0^+$, the variation of B reads [6, 7]

$$\begin{aligned}\left\langle \Delta \widehat{B}(t) \right\rangle &= F \int_{-\infty}^{0} \phi_{BA}(t - t')\, dt' \\ &= F \int_{t}^{\infty} \phi_{BA}(t')\, dt' \equiv F \Phi_{BA}(t),\end{aligned} \tag{1.29}$$

where Φ_{AB} is the *relaxation* function that describes the evolution of the system towards equilibrium after the perturbation has been removed [6]. Thanks to Eq. (1.28), also the relaxation function can be evaluated as an ensemble average,

$$\Phi_{AB}(t) = \beta \left(\left\langle \widehat{A} \widehat{B}(t) \right\rangle_0 - \left\langle \widehat{A} \right\rangle_0 \left\langle \widehat{B} \right\rangle_0 \right). \tag{1.30}$$

In order for Eq. (1.30) to hold, the system must be ergodic, and therefore a generic observable A must satisfy [6, 7]

$$\begin{aligned}\left\langle \widehat{A} \right\rangle_0 &= \lim_{t \to \infty} \frac{1}{t} \int_0^t \widehat{A}(t)\, dt , \\ \lim_{t \to \infty} \left\langle \widehat{A} \widehat{A}(t) \right\rangle_0 &= \left\langle \widehat{A} \right\rangle_0^2 .\end{aligned} \tag{1.31}$$

Quantum formulation of the Green-Kubo response theory

The quantum mechanical formulation of the Green Kubo (GK) theory is similar to the classical one [6, 7]. The density matrix, $\widehat{\rho}$, is a constant of motion, i.e. $\left[\widehat{\mathcal{H}}, \widehat{\rho}\right] = 0$. The force acting on the system perturbs the density matrix, that at first order, $\widehat{\rho}(t) = \widehat{\rho} + \widehat{\rho}\,'(t)$, evolves as [6, 7]

$$\frac{d}{dt} \widehat{\rho}(t) = \frac{1}{i\hbar} \left[\widehat{\mathcal{H}} + \widehat{\mathcal{H}}'(t), \widehat{\rho}(t) \right], \tag{1.32}$$

with the initial condition $\widehat{\rho}\,'(-\infty) = \widehat{\rho}$. The solution of Eq. (1.32) for $\widehat{\rho}\,'(t)$ is

$$\widehat{\rho}\,'(t) = -\frac{1}{i\hbar} \int_{-\infty}^{t} \exp\left\{ -\frac{i\,(t - t')}{\hbar} \widehat{\mathcal{H}} \right\} \left[\widehat{A}, \widehat{\rho} \right] \exp\left\{ \frac{i\,(t - t')}{\hbar} \widehat{\mathcal{H}} \right\} F(t')\, dt' . \tag{1.33}$$

As in classical GK theory, the variation of $\widehat{B}$ due to the perturbed distribution function reads [6, 7]

$$\left\langle \Delta \widehat{B}(t) \right\rangle \equiv \mathrm{Tr}\left\{ \widehat{\rho}\,'(t) \widehat{B} \right\} \tag{1.34}$$

$$= -\frac{1}{i\hbar} \mathrm{Tr}\left\{ \int_{-\infty}^{t} \left[\widehat{A}, \widehat{\rho} \right] \widehat{B}(t - t') F(t')\, dt' \right\}. \tag{1.35}$$

In the Heisenberg picture the operators evolve in time as $\dot{\widehat{B}}(t) = \frac{1}{i\hbar}\left[\widehat{B}(t), \widehat{\mathcal{H}}\right]$, with the initial condition $\widehat{B}(0) = \widehat{B}$. The quantum response function is [6, 7]

$$\phi_{BA}(t) = -\frac{1}{i\hbar}\,\mathrm{Tr}\left\{\left[\widehat{A}, \widehat{\rho}\right]\widehat{B}(t)\right\}, \tag{1.36}$$

which under the assumption that the ensemble is canonical becomes

$$\begin{aligned}\phi_{BA}(t) &= \int_0^\beta \mathrm{Tr}\left\{\widehat{\rho}\dot{\widehat{A}}(i\hbar\lambda)\widehat{B}(t)\,d\lambda\right\} \\ &= -\int_0^\beta \mathrm{Tr}\left\{\widehat{\rho}\widehat{A}(i\hbar\lambda)\dot{\widehat{B}}(t)\,d\lambda\right\}.\end{aligned} \tag{1.37}$$

In the proof of Eq. (1.37), I have used the useful relation [6, 7]

$$\begin{aligned}\left[\widehat{A}, \exp\left\{-\beta\widehat{\mathcal{H}}\right\}\right] &= \exp\left\{-\beta\widehat{\mathcal{H}}\right\}\left(\exp\left\{\beta\widehat{\mathcal{H}}\right\}\widehat{A}\exp\left\{-\beta\widehat{\mathcal{H}}\right\} - \widehat{A}\right) \\ &= \exp\left\{-\beta\widehat{\mathcal{H}}\right\}\int_0^\beta \exp\left\{\lambda\widehat{\mathcal{H}}\right\}\left[\widehat{A}, \widehat{\mathcal{H}}\right]\exp\left\{-\lambda\widehat{\mathcal{H}}\right\}d\lambda\,.\end{aligned} \tag{1.38}$$

Eqs. (1.28) (1.36) are the key elements for the computation of responses to dynamical perturbations. Since the extension to the quantum case can be easily derived, in the following I address only classical systems.

Classical Green-Kubo response theory to a periodic perturbation

I want to apply the classical GK theory to investigate transport phenomena. Eq. (1.29) relates the relaxation function and the response function,

$$\begin{aligned}\phi_{BA}(t) &= -\dot{\Phi}_{BA}(t) \\ &= -\Phi_{\dot{B}A}(t) \\ &= \Phi_{B\dot{A}}(t),\end{aligned} \tag{1.39}$$

$$\Phi_{BA}(-t) = \tau_A\tau_B\Phi_{BA}(t), \tag{1.40}$$

where τ_A is the parity of the observable A under time-reversal symmetry [6, 7]. Complementarily to the adiabatic approach, in which an external agent switches on the perturbation at $t = -\infty$ and removes it at $t = 0$, one can imagine a different path where the system is at equilibrium at $t = 0^-$ and then a constant perturbation is applied. Under the hypothesis that the relaxation from a non-equilibrium condition is indistinguishable to the time evolution of a random fluctuation at equilibrium, the two *gedankenexperiments* are equivalent.

In space translational invariant systems, perturbations at finite wavelength are coupled to densities of the same wave-vector, compatible with the enforced Periodic Boundary Conditions (PBC), as in:

$$\begin{aligned}\Phi_{ba}(\mathbf{r}, t; \mathbf{r}', t') &= -V\beta\left\langle\left(\widehat{b}(\mathbf{r}, t) - \left\langle\widehat{b}(\mathbf{r})\right\rangle_0\right)\left(\widehat{a}(\mathbf{r}', t') - \left\langle\widehat{a}(\mathbf{r}')\right\rangle_0\right)\right\rangle_0 \\ &= \int\frac{d\mathbf{k}}{(2\pi)^3}\int\frac{d\omega}{2\pi}\Phi_{ba}(\mathbf{k}, \omega)e^{-i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')+i\omega(t-t')}.\end{aligned} \tag{1.41}$$

The Fourier transform of the relaxation function [13] is connected to the Fourier transform of the response function, χ_{ba}, the so-called *admittance* [13],

$$2i\phi_{ba}(\mathbf{r},t;\mathbf{r}',t')\Theta(t-t') = \int \frac{d\mathbf{k}}{(2\pi)^3}\int \frac{d\omega}{2\pi}\chi_{ba}(\mathbf{k},\omega)e^{-i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')+i\omega(t-t')}. \quad (1.42)$$

The admittance is a complex-valued function [13],

$$\begin{aligned}\chi_{ba}(\mathbf{k},\omega) &= \chi'_{ba}(\mathbf{k},\omega) + i\chi''_{ba}(\mathbf{k},\omega)\\ &= \lim_{\epsilon\to 0^+}\chi_{ba}(\mathbf{k},\omega+i\epsilon),\end{aligned} \quad (1.43)$$

and its imaginary part satisfies the fluctuation dissipation theorem [13, 14],

$$\frac{1}{2}\Phi_{ba}(\mathbf{k},\omega) = \frac{\chi''_{ba}(\mathbf{k},\omega)}{\omega}. \quad (1.44)$$

Let us consider an equilibrium system perturbed at $t = 0$ by a potential, $v_a(\mathbf{r},t)$, space and time dependent: $\widehat{\mathcal{H}'} = Vv_a(\mathbf{k})\Theta(t)\hat{a}(-\mathbf{k},t)$, where $v_a(\mathbf{k})$ and $\hat{a}(\mathbf{k})$ are the space Fourier components as in $f(\mathbf{k}) = \frac{1}{V}\int_V f(\mathbf{r})e^{-i\mathbf{k}\cdot\mathbf{r}}\,d\mathbf{r}$. The perturbation induces a flow of the conserved density b, i. e. a current $\boldsymbol{j}_b(\mathbf{k},t)$, that can be computed from the GK theory as [15, 16]:

$$\begin{aligned}\left\langle\hat{\boldsymbol{j}}_b(\mathbf{k},t)\right\rangle &= \Phi_{\boldsymbol{j}_b a}(\mathbf{k},t)v_a(\mathbf{k})\\ \Phi_{\boldsymbol{j}_b a}(\mathbf{k},t) &= -V\beta\left\langle\hat{\boldsymbol{j}}_b(\mathbf{k},t)\hat{a}(-\mathbf{k})\right\rangle_0,\end{aligned} \quad (1.45)$$

where $\boldsymbol{J}_b(t) = \frac{1}{V}\int_V \boldsymbol{j}_b(\mathbf{r},t)$. The currents and the densities associated to conserved quantities obey the continuity equations:

$$\begin{aligned}-i\mathbf{k}\cdot\hat{\boldsymbol{j}}_a(\mathbf{k},t) &= \dot{\hat{a}}(\mathbf{k},t),\\ -i\mathbf{k}\cdot\hat{\boldsymbol{j}}_b(\mathbf{k},t) &= \dot{\hat{b}}(\mathbf{k},t).\end{aligned} \quad (1.46)$$

According to the GK theory, the relaxation function can be computed as

$$\begin{aligned}\phi_{\boldsymbol{j}_b a}(\mathbf{k},t) &= -V\beta\left\langle\hat{\boldsymbol{j}}_b(\mathbf{k},t)\dot{\hat{a}}(-\mathbf{k})\right\rangle_0 = -i\mathbf{k}V\beta\left\langle\hat{\boldsymbol{j}}_b(\mathbf{k},t)\hat{\boldsymbol{j}}_a(-\mathbf{k})\right\rangle_0\\ &= \frac{i\mathbf{k}}{k^2}V\beta\left\langle\dot{\hat{b}}(\mathbf{k},t)\dot{\hat{a}}(-\mathbf{k})\right\rangle_0 = -\frac{i\mathbf{k}}{k^2}\ddot{\Phi}_{ba}(\mathbf{k},t),\end{aligned} \quad (1.47)$$

$$\begin{aligned}\Phi_{\boldsymbol{j}_b a}(\mathbf{k},t) &= \int_0^t \phi_{\boldsymbol{j}_b a}(\mathbf{k},t')\,dt'\\ &= -\frac{i\mathbf{k}}{k^2}\dot{\Phi}_{ba}(\mathbf{k},t).\end{aligned} \quad (1.48)$$

Eq. (1.45) can be rewritten in the form

$$\begin{aligned}\left\langle\hat{\boldsymbol{j}}_b(\mathbf{k},t)\right\rangle &= \left(-V\beta\int_0^t\left\langle\hat{\boldsymbol{j}}_b(\mathbf{k},t')\dot{\hat{a}}(-\mathbf{k})\right\rangle_0 dt'\right)v_a(\mathbf{k})\\ &= \left(V\beta\int_0^t\left\langle\hat{\boldsymbol{j}}_b(\mathbf{k},t')\hat{\boldsymbol{j}}_a(-\mathbf{k})\right\rangle_0 dt'\right)\cdot(-i\mathbf{k}v_a(\mathbf{k})).\end{aligned} \quad (1.49)$$

In the homogeneous limit, $\mathbf{k} \to 0$, I evaluate the time dependent response of the current to the thermodynamic force, i.e. the gradient of the potential.

$$\left\langle \widehat{\boldsymbol{J}}_b \right\rangle = \lim_{t\to\infty} \left(-\frac{V\beta}{3} \int_0^t \lim_{\mathbf{k}\to 0} \left\langle \widehat{\boldsymbol{j}}_b(\mathbf{k}, t') \cdot \widehat{\boldsymbol{j}}_a(-\mathbf{k}) \right\rangle_0 dt' \right) \nabla v_a. \tag{1.50}$$

The macroscopic homogeneous response to the gradient of the potential is

$$\left\langle \widehat{\boldsymbol{J}}_b \right\rangle = \frac{V\beta}{3} \int_0^\infty \left\langle \widehat{\boldsymbol{J}}_b(t) \cdot \widehat{\boldsymbol{J}}_a \right\rangle_0 dt \, \nabla v_a. \tag{1.51}$$

Therefore, thanks to the classical GK theory of linear response, the Onsager coefficients can be computed from the equilibrium fluctuations of the currents:

$$L_{ab} = \frac{V\beta}{3} \int_0^\infty \left\langle \widehat{\boldsymbol{J}}_b(t) \cdot \widehat{\boldsymbol{J}}_a \right\rangle_0 dt \, . \tag{1.52}$$

The Helfand-Einstein formulation of transport coefficients

Each Onsager coefficient can be evaluated as a GK integrals, defined in Eq. (1.52). Thanks to the seminal work of Einstein [17], that relates the mean squared displacement of a diffusing particle to its velocity autocorrelation function, Helfand discussed an alternative expression for the transport coefficients [18]. For a generic stationary stochastic process, $\widehat{X}_t$, one has [15]:

$$\frac{1}{\tau} \left\langle \left| \int_0^\tau \widehat{X}_t \, dt \right|^2 \right\rangle_0 = 2 \int_0^\tau \left\langle \widehat{X}_t \widehat{X}_0 \right\rangle_0 dt - \frac{2}{\tau} \int_0^\tau \left\langle \widehat{X}_t \widehat{X}_0 \right\rangle_0 t \, dt \, , \tag{1.53}$$

where, in the $\tau \to \infty$ limit, the second term of the right-hand side is negligible. When the generic stochastic process is a flux, Eq. (1.53) provides a new, equivalent with respect to GK integrals, method for computing Onsager coefficients:

$$L_{aa} = \lim_{\tau\to\infty} \frac{V\beta}{3} \int_0^\tau \left(1 - \frac{t}{\tau} \right) \left\langle \widehat{\boldsymbol{J}}_a(t) \cdot \widehat{\boldsymbol{J}}_a \right\rangle_0 dt \tag{1.54}$$

$$= \lim_{\tau\to\infty} \frac{V\beta}{6} \left\langle \left| \int_0^\tau \widehat{\boldsymbol{J}}_a(t) \right|^2 \right\rangle_0 dt \, . \tag{1.55}$$

The HE formula is an unbiased estimator of the Onsager coefficient and is more statistically efficient than the GK integral [19].

The Fick's law

The order of the limits in Eq. (1.50) is crucial. At finite $\mathbf{k}$, in the limit $t \to \infty$, the system is at equilibrium with the perturbation $v_a(\mathbf{k})\Theta(t)$ and the variations of the perturbed densities are zero: $\left\langle \hat{b}(\mathbf{k}, t \to \infty) \right\rangle = \left\langle \hat{a}(\mathbf{k}, t \to \infty) \right\rangle = 0$. Thanks to the continuity equations, after having imposed a constant perturbation at $t = 0$ the $\mathbf{k}$-Fourier components of the currents vanish at equilibrium. When I invert the order of the limits in Eq. (1.50), the transport coefficient vanishes as well. In the long-time limit, when the fluctuations of the densities

become uncorrelated, Eq. (1.48), that is the finite $\mathbf{k}$ estimator of the transport coefficient, is zero [16]. On the contrary, the correlation of the macroscopic fluxes in Eq. (1.51) (1.52) are not affected by the same challenges. In the linear regime and assuming that $b(\mathbf{r}, t)$ is the only relevant conserved density, the hydrodynamic Fick's law holds:

$$\begin{aligned} \left\langle \widehat{j}_b(\mathbf{r}, t) \right\rangle &= -D_b \nabla \left\langle \widehat{b}(\mathbf{r}, t) \right\rangle \\ \left\langle \dot{\widehat{b}}(\mathbf{r}, t) \right\rangle &= D_b \nabla^2 \left\langle \widehat{b}(\mathbf{r}, t) \right\rangle , \end{aligned} \tag{1.56}$$

where D_b is a diffusion constant. After a brief inspection of Eq. (1.56), the hydrodynamic time scale $\frac{1}{D_b k^2}$ emerges. In this period of time, the initial inhomogeneity is reduced, but not vanishes. Furthermore, the current has reached a plateau before the fluctuation relaxes towards thermodynamic equilibrium [3]: see Fig.2 in Ref. [16]. In setting $\mathbf{k}$ equal to zero, the dynamics of the fluctuation is locked in the hydrodynamic regime [16, 20],

$$\begin{aligned} \left\langle \widehat{b}(\mathbf{k}, t) \right\rangle &= \left\langle \widehat{b}(\mathbf{k}, 0) \right\rangle e^{-D_b k^2 t} \\ \left\langle \widehat{j}(\mathbf{k}, t) \right\rangle &= -D_b \left(\mathbf{k} \left\langle \widehat{b}(\mathbf{k}, 0) \right\rangle e^{-D_b k^2 t} \right) . \end{aligned} \tag{1.57}$$

An important caveat must be stressed: in systems with long-range interaction forces, the correlation function is not equal to the transport coefficient even in the long-wavelength limit [11, 13], see for details Appendix A and Appendix B.

1.2.2 *Representation of the heat flux*

In the microcanonical ensemble the rate of production of entropy density, $\dot{s}$, is [21]

$$\dot{s} = \boldsymbol{J}_e \cdot \nabla \frac{1}{T} + \sum_{k=1}^{n} \boldsymbol{J}_k \cdot \nabla \left(-\frac{\mu_k}{T} \right) , \tag{1.58}$$

where $\boldsymbol{J}_e$ is the energy flux, n is the number of species and $\boldsymbol{J}_k$ is the number current of the species k. Through thermodynamic manipulations, one has [21]

$$T \, d\frac{\mu_k}{T} = d\mu_k|_T - \frac{h_k}{T} \, dT . \tag{1.59}$$

In Eq. (1.59) the suffix T indicates that the differential of the chemical potential has been taken at constant temperature and h_ks are the molar partial enthalpies [22–24]. The isothermal gradient of the chemical potential is a thermodynamic force. The *heat flux*, i.e. the current associated to the gradient of the temperature field, is [21]

$$\begin{aligned} \boldsymbol{J}_q &= \boldsymbol{J}_e - \sum_{k=1}^{n} h_k \boldsymbol{J}_k \\ \dot{s} &= -\frac{1}{T^2} \boldsymbol{J}_q \cdot \nabla T - \frac{1}{T} \sum_{k=1}^{n} \boldsymbol{J}_k \cdot \nabla \mu_k|_T . \end{aligned} \tag{1.60}$$

Equivalently, the derivation of heat density follows from the hypothesis of local thermal equilibrium considering the interactions with the surrounding environment. Let the system be partitioned in many blocks which allow energy and particle exchange. In local thermal equilibrium, the Boltzmann factor is:

$$\exp\left\{-\int_V \frac{\widehat{e}(\mathbf{r}) - \sum_k \mu_k(\mathbf{r})\widehat{n}_k(\mathbf{r})}{T(\mathbf{r})}\,d\mathbf{r}\right\}. \tag{1.61}$$

I am assuming that the chemical potential $\mu(\mathbf{r})$ depends on $\mathbf{r}$ only through $T(\mathbf{r})$. Thanks to Eq. 1.59, assuming that $T(\mathbf{r}) = T + T'(\mathbf{r})$ and expanding at linear order in $T'(\mathbf{r})$, one can write:

$$\int_V \frac{\widehat{e}(\mathbf{r}) - \sum_k \mu_k\left(T(\mathbf{r})\right)\widehat{n}_k(\mathbf{r})}{T(\mathbf{r})}\,d\mathbf{r} \approx \frac{\widehat{E} - \sum_k \mu_k \widehat{N}_k}{T} - \int_V T'(\mathbf{r})\frac{\widehat{e}(\mathbf{r}) - h_k\widehat{n}_k(\mathbf{r})}{T^2}\,d\mathbf{r}\,. \tag{1.62}$$

where $\widehat{e}(\mathbf{r}) - h_k\widehat{n}_k(\mathbf{r})$ is the heat density, which is a mechanical quantity that can be evaluated from a equilibrium molecular dynamics simulation. The definition of the thermodynamic force and of the flux coupled to the temperature gradient are interdependent. Ref.[21] explains that each definition describes a different experimental setup. Of course, this arbitrariness leaves all physical results unchanged. Nevertheless, one specific choice might result in more transparent Onsager coefficients, which can be directly compared to the transport coefficients entering the hydrodynamic equations.

Mori's approach in the grand canonical ensemble

As expressed above, ensembles play a central role in the definition of the Onsager coefficients. In single component systems, the microcanonical average of the energy and the heat flux autocorrelation coincide. Transport theory has been derived both in the microcanonical ensemble, as proposed by Kubo and Green [6, 9], and in the grand canonical ensemble [25], as developed by Mori in a groundbreaking paper in 1958 [25]. In Mori's approach, the heat flux Onsager coefficient is computed as

$$L_{qq} = \frac{V}{3k_B T^2}\int_0^\infty \left\langle \widehat{\boldsymbol{J}}_q(t)\cdot\widehat{\boldsymbol{J}}_q\right\rangle_{g.c.} dt\,, \tag{1.63}$$

where $\langle\circ\rangle_{g.c.}$ is the expectation value over the grand canonical ensemble. Mori pointed out that, even in single component systems, the enthalpic term of the heat flux in L_{qq} yields a non-vanishing contribution, apparently contradicting the work of Green and Kubo which relies on the microcanonical ensemble averages. The conundrum was solved separately by Green and Martin in the early 1960s [20, 26]. Green's comment focuses on relaxation of fluctuations in different ensembles. In the grand canonical ensemble, the relaxation of the fluctuations describes the reach of macroscopic equilibrium, whereas in the microcanonical ensemble, is related to the microscopic processes of the fluid [26]. Martin discussed the impact of computing the autocorrelation of the

macroscopic fluxes setting $\mathbf{k}$ equal to 0, and realized that these homogeneous correlations can actually depend on the ensemble. The hydrodynamic equations themselves are ensemble dependent because they are strongly sensitive to the initial and boundary conditions. Nevertheless, in taking the $\mathbf{k} \to 0$ limit in Eq (1.50), and making sure that $\frac{1}{k}$ is much smaller than the linear dimension of the system, these complications become negligible [20]. In the $V \to \infty$ and *then* the $\mathbf{k} \to 0$ limit, the choice of ensemble and flux is irrelevant [20]. Otherwise, considering the non-physical limit, $\mathbf{k} = 0$ at finite V, there exist a specific relation between fluxes and ensembles [20]. The fluxes and ensembles considered in Green and Kubo's, and Mori's work, result in the correct evaluation of the nonphysical limit [20, 26].

1.3 INVARIANCE PRINCIPLES IN THE THEORY OF TRANSPORT

1.3.1 *Gauge invariance principle of transport coefficients*

The definition of the density, $b(\mathbf{r})$ depends on the partitioning of the extensive quantity $B = \int_V b(\mathbf{r})\, d\mathbf{r}$ into local contributions. Two densities $b(\mathbf{r})$ and $b'(\mathbf{r})$ are equivalent if their integral over the volume coincides: $\int_V b(\mathbf{r})\, d\mathbf{r} = \int_V b'(\mathbf{r})\, d\mathbf{r}$. Two equivalent densities differ only by the divergence of a bounded vector, $\boldsymbol{p}(\mathbf{r})$,

$$b'(\mathbf{r}, t) = b(\mathbf{r}, t) - \nabla \cdot \boldsymbol{p}(\mathbf{r}, t). \tag{1.64}$$

Each density is linked to its respective current, $\boldsymbol{j}_b(\mathbf{r}, t)$ and $\boldsymbol{j}'_b(\mathbf{r}, t)$, by the continuity equation:

$$\begin{aligned} \boldsymbol{j}'_b(\mathbf{r}, t) &= \boldsymbol{j}_b(\mathbf{r}, t) + \dot{\boldsymbol{p}}(\mathbf{r}, t) \\ \boldsymbol{J}'_b(t) &= \boldsymbol{J}_b(t) + \dot{\boldsymbol{P}}(t), \end{aligned} \tag{1.65}$$

where $\boldsymbol{P}(t) = \frac{1}{V}\int_V \boldsymbol{p}(\mathbf{r}, t)\, d\mathbf{r}$. The two definitions of the density lead to two currents, that differ by the time derivative of a bounded vector. For energy density, the bounded vector $\boldsymbol{p}(\mathbf{r}, t)$ reflects the arbitrariness on partitioning of the total energy into local contributions [27]. Thanks to the Helfand-Einstein formulation of Onsager coefficients, Eqs. (1.55) (1.55), the coefficient associated to the current $\boldsymbol{J}'_b(t)$, L'_{bb}, can be written as:

$$L'_{bb} = c \lim_{\tau\to\infty} \frac{1}{2\tau} \left\langle \left| \int_0^\tau \widehat{\boldsymbol{J}}'_b(t)\, dt \right|^2 \right\rangle_0 \tag{1.66}$$

where c is the appropriate constant that depends on the flux being considered and

$$\boldsymbol{D}_b(\tau) = \int_0^\tau \boldsymbol{J}_b(t)\, dt \tag{1.67}$$

is the generalized dipole associated to the current $\boldsymbol{J}_b$. The Onsager coefficient L'_{BB} can be computed as

$$L'_{bb} = c \lim_{\tau\to\infty} \frac{\left\langle \left|\widehat{\boldsymbol{D}}_b(\tau)\right|^2 \right\rangle_0 + 2\left\langle \widehat{\boldsymbol{D}}_b(\tau)\cdot\widehat{\boldsymbol{P}}(\tau)\right\rangle_0 + \left\langle \left|\widehat{\boldsymbol{P}}(\tau)\right|^2 \right\rangle_0}{2\tau}. \tag{1.68}$$

On one hand, as anticipated by Einstein in 1905 [17] and clarified by Helfand in 1960 [18], the variance of the generalized dipole of a diffusive process grows linearly in time. On the other, the variance of $\boldsymbol{P}(\tau)$ is limited by definition:

$$\begin{aligned} \left\langle \left|\widehat{\boldsymbol{D}}_b(\tau)\right|^2 \right\rangle_0 &\sim \mathcal{O}(\tau) \\ \left\langle \widehat{\boldsymbol{D}}_b(\tau)\cdot\widehat{\boldsymbol{P}}(\tau)\right\rangle_0 &\sim \mathcal{O}(\tau^{1/2}) \\ \left\langle \left|\widehat{\boldsymbol{P}}_b(\tau)\right|^2 \right\rangle_0 &\sim \mathcal{O}(1). \end{aligned} \tag{1.69}$$

In the long time limit, thanks to Eqs. (1.68) and (1.69), the transport coefficient associated with the modified current $\boldsymbol{J}'_b$ coincides with the one derived from $\boldsymbol{J}_b$,

$$L'_{BB} = c \lim_{\tau\to\infty} \frac{\left\langle \left|\widehat{\boldsymbol{D}}_b(\tau)\right|^2 \right\rangle_0}{2\tau} = L_{BB}. \tag{1.70}$$

Eq. (1.70) synthesizes the so-called *gauge invariance principle* of transport coefficients [15, 19]. The addition of the divergence of a bounded vector field changes the density and also the resulting macroscopic current, nevertheless it does not affect the measurable Onsager coefficient. The invariance principle of the Onsager coefficients has only been studied recently but it has been applied for the computation of thermal conductivity, electric conductivity and oxidation charges [19, 28–36].

1.3.2 *Convective invariance of heat conductivity in multicomponent systems*

Thermal conductivity is defined in absence of convective flows as

$$\kappa \equiv \left.\frac{\left\langle \widehat{J}_e \right\rangle}{\nabla T}\right|_{J_k=0}. \tag{1.71}$$

In Eq. (1.71), I consider any Cartesian component of the energy current and temperature gradient. To address heat transport, the expression of the energy flux, even if it depends on the specific decomposition of the energy density into local contributions, is crucial. It can be formally derived from the energy density:

$$\begin{aligned} e(\mathbf{r},\Gamma) &= \sum_n \delta(\mathbf{r}-\mathbf{R}_n) e_n(\Gamma) \\ e_n(\Gamma) &= \frac{\mathbf{P}_n^2}{2M_n} + v_n\left(\{\mathbf{R}\}\right), \end{aligned} \tag{1.72}$$

where $\mathbf{R}_n$ is the position of the n-th atom, $\mathbf{P}_n$ is the canonical momentum and $v_n(\{\mathbf{R}\})$ is the per-atom decomposition of the potential energy, $\mathcal{V}$, such that $\sum_n v_n = \mathcal{V}$. $\boldsymbol{J}_e$ is evaluated via the continuity equation as

$$\begin{aligned}\boldsymbol{J}_e(\Gamma) &= \frac{1}{V}\int_V \dot{e}(\mathbf{r},\Gamma)\mathbf{r}\,d\mathbf{r}\\ &= \frac{1}{V}\int_V \left[\sum_n \left(\frac{\partial e(\mathbf{r},\Gamma)}{\partial \mathbf{R}_n}\cdot\mathbf{V}_n + \frac{\partial e(\mathbf{r},\Gamma)}{\partial \mathbf{P}_n}\cdot\mathbf{F}_n\right)\right]\mathbf{r}\,d\mathbf{r}\\ &= \frac{1}{V}\left[\sum_n e_n\mathbf{V}_n + \sum_{n,m}(\mathbf{R}_n-\mathbf{R}_m)\,\mathbf{F}_{nm}\cdot\mathbf{V}_n\right],\end{aligned} \tag{1.73}$$

where $\mathbf{V}_n = \mathbf{P}_n/M_n$ and $\mathbf{F}_{nm} = -\frac{\partial v_n}{\partial \mathbf{R}_m}$ is the n-th contribution to the force acting on the m-th atom.

In single component systems, κ is proportional to the energy Onsager coefficient:

$$L_{ee} = \frac{V}{3k_B T}\int_0^\infty \left\langle \widehat{\boldsymbol{J}}_e(t)\cdot\widehat{\boldsymbol{J}}_e(0)\right\rangle_0 dt\,. \tag{1.74}$$

In a two-component system, the Onsager equations read

$$\begin{aligned}\left\langle\widehat{\boldsymbol{J}}_e\right\rangle &= L_{ee}\nabla\frac{1}{T} + L_{e1}\nabla\left(\frac{-\mu_1}{T}\right) + L_{e2}\nabla\left(\frac{-\mu_2}{T}\right)\\ \left\langle\widehat{\boldsymbol{J}}_1\right\rangle &= L_{1e}\nabla\frac{1}{T} + L_{11}\nabla\left(\frac{-\mu_1}{T}\right) + L_{12}\nabla\left(\frac{-\mu_2}{T}\right)\\ \left\langle\widehat{\boldsymbol{J}}_2\right\rangle &= L_{2e}\nabla\frac{1}{T} + L_{21}\nabla\left(\frac{-\mu_1}{T}\right) + L_{22}\nabla\left(\frac{-\mu_2}{T}\right).\end{aligned} \tag{1.75}$$

Thermal conductivity, defined in Eq. (1.71), in a M-component system is

$$\kappa = \frac{1}{T}\left[L_{ee} - \sum_{k,p=1}^{M-1} L_{ke}\left(\mathbf{L}^{-1}\right)_{kp} L_{ep}\right], \tag{1.76}$$

where

$$L_{ke} = \frac{V}{3k_B T}\int_0^\infty \left\langle \widehat{\boldsymbol{J}}_k(t)\cdot\widehat{\boldsymbol{J}}_e(0)\right\rangle_0 dt\,, \tag{1.77}$$

$$L_{kp} = \frac{V}{3k_B T}\int_0^\infty \left\langle \widehat{\boldsymbol{J}}_k(t)\cdot\widehat{\boldsymbol{J}}_p(0)\right\rangle_0 dt\,. \tag{1.78}$$

Notably, Eq. (1.76) is invariant under the transformation $\boldsymbol{J}_e \to \boldsymbol{J}_e + \sum_{k=1}^{M} c_k\boldsymbol{J}_k$, and c_k can be any real number. This invariance principle is dubbed *convective invariance* of thermal conductivity and is based on the hydrodynamic definition of the transport coefficient.

1.4 HYDRODYNAMIC DESCRIPTION OF A DIFFUSIVE PROCESS

The hydrodynamic description of transport presented in Ref. [14] is very powerful. Transport in a normal —single component and isotropic— fluid is described by the conservation laws, constitutive equations and thermodynamic

relations [14, 20]. In such systems the number of particles, energy and momentum are conserved. The continuity equations, that hold microscopically, are:

$$\begin{aligned}
&\partial_t n(\mathbf{r},t) + \nabla\cdot\frac{\boldsymbol{j}(\mathbf{r},t)}{m} = 0\\
&\partial_t \boldsymbol{j}(\mathbf{r},t) + \nabla\cdot\boldsymbol{\tau}(\mathbf{r},t) = 0\\
&\partial_t e(\mathbf{r},t) + \nabla\cdot\boldsymbol{j}_e(\mathbf{r},t) = 0,
\end{aligned} \tag{1.79}$$

where m is the particle mass, n is the number density, e is the energy density, $\boldsymbol{j}$ is the momentum density and $\boldsymbol{\tau}$ is the stress tensor that serves as a momentum current. The continuity equations are not sufficient to describe the dynamics of the fluid. Under the assumptions that the fluctuations of the densities in space and time are "slow" [14, 21], and that the system is in local thermodynamic equilibrium, the *constitutive* equations hold. They link the five conserved quantities, which characterize completely the state of the fluid, with the intensive variables, that are conventionally chosen to be temperature, $T(\mathbf{r},t)$, pressure, $p(\mathbf{r},t)$, and, average velocity, $\mathbf{v}(\mathbf{r},t)$ [14]:

$$\begin{aligned}
\left\langle\widehat{\boldsymbol{j}}(\mathbf{r},t)\right\rangle &= \langle\widehat{n}(\mathbf{r},t)\rangle\, m\mathbf{v}(\mathbf{r},t) = nm\mathbf{v}(\mathbf{r},t)\\
\langle\widehat{\tau}_{\alpha\beta}(\mathbf{r},t)\rangle &= p(\mathbf{r},t)\delta_{\alpha\beta} - \eta\left[\nabla_\alpha v_\beta(\mathbf{r},t) + \nabla_\beta v_\alpha(\mathbf{r},t) - \frac{2}{3}\nabla\cdot\mathbf{v}(\mathbf{r},t)\delta_{\alpha\beta}\right]\\
&\quad - \xi\nabla\cdot\mathbf{v}(\mathbf{r},t)\delta_{\alpha\beta}\\
\left\langle\widehat{\boldsymbol{j}}_e(\mathbf{r},t)\right\rangle &= (\langle\widehat{e}\rangle + p)\mathbf{v}(\mathbf{r},t) - \kappa\nabla T(\mathbf{r},t).
\end{aligned} \tag{1.80}$$

Eqs. (1.80) are the linearized Navier-Stokes equations and η, ξ and κ are shear and bulk viscosity, and thermal conductivity respectively [14].

Combining Eq. (1.80) and Eq. (1.79), Ref. [14] provides a detailed derivation of the hydrodynamic functional form of the densities response functions in terms of thermodynamic derivatives and transport coefficients. In the long-wavelength limit, the static susceptibilities are equal to some thermodynamic derivatives. For a generic quantity A, one has

$$\left.\frac{\partial A}{\partial\mu}\right|_{T,V} = \int\frac{d\omega}{\pi}\frac{\chi''_{An}(\mathbf{k}=0,\omega)}{\omega}. \tag{1.81}$$

Thanks to Eq. (1.81), one derives the sum rules [14, 20, 37]

$$\begin{aligned}
\lim_{\mathbf{k}=0}\int\frac{d\omega}{\pi}\frac{\chi''_{nn}(\mathbf{k},\omega)}{\omega} &= \lim_{\mathbf{k}=0}\chi_{nn}(\mathbf{k}) = n\left.\frac{\partial n}{\partial p}\right|_T\\
\lim_{\mathbf{k}=0}\int\frac{d\omega}{\pi}\frac{\chi''_{nq}(\mathbf{k},\omega)}{\omega} &= \lim_{\mathbf{k}=0}\chi_{nq}(\mathbf{k}) = T\left.\frac{\partial n}{\partial T}\right|_p\\
\lim_{\mathbf{k}=0}\int\frac{d\omega}{\pi}\frac{\chi''_{qq}(\mathbf{k},\omega)}{\omega} &= \lim_{\mathbf{k}=0}\chi_{qq}(\mathbf{k}) = nmc_pT.
\end{aligned} \tag{1.82}$$

where $q = e - hn$ is the heat density as in Eq. (1.62), h is the per unit volume enthalpy and $mnc_p = \frac{T}{V}\left.\frac{\partial S}{\partial T}\right|_p$. Notice that, in the canonical ensemble,

where the number of particles is fixed, $\chi'_{nn}(\mathbf{k}=0)=0$ and in the micro-canonical ensemble $\chi'_{qq}(\mathbf{k}=0)=0$. Even if the macroscopic fluctuations of conserved densities are ensemble dependent, $\chi'(\mathbf{k})$ describes local fluctuations that are independent on the boundary conditions [14]. Building on Martin discussion [20], one must consider first the $V\to\infty$ limit, and then $\mathbf{k}\to 0$.

As derived in Ref. [14], the absorptive parts of the densities response functions read

$$\frac{1}{\omega}\chi''_{nn}(\mathbf{k},\omega)=m\left.\frac{\partial n}{\partial p}\right|_T\left[\frac{c_v/c_p c^2k^4\Gamma_B}{(\omega^2-c^2k^2)^2+(\omega k^2\Gamma_B)^2}+\frac{(1-c_v/c_p)k^2D_T}{\omega^2+(k^2D_T)^2}-\left(1-\frac{c_v}{c_p}\right)\frac{(\omega^2-c^2k^2)k^2D_T}{(\omega^2-c^2k^2)^2+(\omega k^2\Gamma_B)^2}\right] \tag{1.83}$$

$$\frac{1}{\omega}\chi''_{qq}(\mathbf{k},\omega)=mnc_pT\frac{k^2D_T}{\omega^2+(k^2D_T)^2} \tag{1.84}$$

$$\frac{1}{\omega}\chi''_{nq}(\mathbf{k},\omega)=T\left.\frac{\partial n}{\partial T}\right|_p\left[\frac{k^2D_T}{\omega^2+(k^2D_T)^2}-\frac{(\omega^2-c^2k^2)k^2D_T}{(\omega^2-c^2k^2)^2+(\omega k^2\Gamma_B)^2}\right], \tag{1.85}$$

where $D_T=\frac{\kappa}{mnc_p}$, $\Gamma_B=D_l+D_T(c_p/c_v-1)$ is the Brillouin peak, $D_l=\frac{4/3\eta+\xi}{mn}$ and $mnc_v=\frac{T}{V}\left.\frac{\partial S}{\partial T}\right|_n$. The bulk viscosity depends on the longitudinal fluctuations of the momentum, Eq. (1.83), that are linked to the numeber density's via the continuity equation:

$$\frac{4}{3}\eta+\xi=\lim_{\omega\to0}\lim_{\mathbf{k}\to0}\frac{m^2\omega^3}{k^4}\chi''_{nn}(\mathbf{k},\omega)=\lim_{\omega\to0}\lim_{\mathbf{k}\to0}\frac{\omega}{k^2}\chi''_l(\mathbf{k},\omega). \tag{1.86}$$

χ_l is the longitudinal component of the momentum response function. Thanks to Eq. (1.84), thermal conductivity of a normal fluid is

$$T\kappa=\lim_{\omega\to0}\lim_{\mathbf{k}\to0}\frac{\omega}{k^2}\chi''_{qq}(\mathbf{k},\omega). \tag{1.87}$$

Interestingly, Eqs. (1.86) (1.87) are Kubo-like formulas.

Since in Eq. (1.86) $\lim_{\mathbf{k}\to0}\frac{1}{k^4}\chi''_{nn}(\mathbf{k},\omega)$ is finite, then $\lim_{\mathbf{k}\to0}\frac{1}{k^2}\chi''_{nn}(\mathbf{k},\omega)$ must vanish. Thanks to these considerations and to Eq. (1.85), I write

$$0=\lim_{\omega\to0}\lim_{\mathbf{k}\to0}\frac{m^2\omega}{k^2}\chi''_{nn}(\mathbf{k},\omega)=\lim_{\omega\to0}\lim_{\mathbf{k}\to0}\frac{m\omega}{k^2}\chi''_{nq}(\mathbf{k},\omega). \tag{1.88}$$

Eqs. (1.88) (1.87) lead to the familiar convective invariance property [14, 20],

$$T\kappa=\lim_{\omega\to0}\lim_{\mathbf{k}\to0}\frac{\omega}{k^2}\chi''_{q+\lambda n,q+\lambda n}(\mathbf{k},\omega)=\lim_{\omega\to0}\lim_{\mathbf{k}\to0}\frac{\omega}{k^2}\chi''_{e,e}(\mathbf{k},\omega),\ \forall\lambda. \tag{1.89}$$

In normal fluids, κ can be computed directly from the fluctuation of the energy density.

2

CEPSTRAL AND BAYESIAN REGRESSION ANALYSIS OF RESPONSE FUNCTIONS

GREEN KUBO linear response theory provides sound theoretical foundations for computing Onsager coefficients via Equilibrium Molecular Dynamics (EMD) simulations. This chapter addresses the numerical challenges associated with GK estimators. First, I analyze the computational drawbacks of the direct time integration of the correlation function of the fluxes. Next, *cepstral analysis* is reviewed, emphasizing its efficiency and limitations. The final sections introduce a Bayesian protocol for the comprehensive estimation of the entire Onsager matrix, along with a Bayesian linear regression scheme designed to extrapolate the long-wavelength limit of periodic response functions.

2.1 STATISTICAL PROPERTIES OF GREEN-KUBO ESTIMATORS

The Onsager coefficient, L_{ab}, can be evaluated from the time-integral of the correlation function of the fluxes, J_a and J_b, as Eq. (1.52) [6–9]:

$$L_{ab} = V\beta \int_0^\infty \left\langle \widehat{J}_b(t)\widehat{J}_a(0) \right\rangle_0 dt\,. \tag{2.1}$$

The GK formula allows to investigate non-equilibrium properties through EMD simulations, offering an alternative to Non Equilibrium Molecular Dynamics (NEMD) methods, whose statistical uncertainty is challenging to derive [38–43]. The Helfand Einstein (HE) formula, Eq. (1.54), reads [18]:

$$L_{ab} = \lim_{\tau\to\infty} V\beta \int_0^\tau \left(1 - \frac{t}{\tau}\right) \left\langle \widehat{J}_a(t)\widehat{J}_b(0) \right\rangle_0 dt\,. \tag{2.2}$$

In this chapter I study statistical techniques aiming to enhance the numerical efficiency of Eqs. (2.1) (2.2) and to provide a sound estimate of the statistical uncertainty [31, 44–46].

In microcanonical simulations, the Hamilton equations of motion are integrated via the velocity-Verlet algorithm, with a finite time-step ϵ. The simulated trajectory itself is a sample of a stochastic process. I denote with the hat, as in $\widehat{\circ}$, functions of the phase space, stochastic processes and their samples. In a EMD trajectory of length $\tau_{tot} \equiv N\epsilon$, I harvest the currents' time-series, denoted as $\{\widehat{J}_a^m = \widehat{J}_a(m\epsilon)\}$. The time-correlation of the fluxes is

$$\widehat{C}_{ab}^m = \frac{1}{N-m} \sum_{l=0}^{N-1-m} \widehat{J}_a^{l+m}\widehat{J}_b^l. \tag{2.3}$$

Thanks to the central limit theorem, $\widehat{C}_{ab}^m$ is a sample of a normal-distributed stochastic process whose variance decreases with N.

The GK and HE estimators of the Onsager coefficient, Eqs. (2.1) (2.2), are

$$L^{X}_{ab}(\tau) = V\beta \int_0^\infty \left\langle \widehat{J}_a(t)\widehat{J}_b(0)\right\rangle_0 \Theta^{\tau}_{X}(t)\,dt\,, \tag{2.4}$$

where X = GK or HE, and

$$\Theta^{\tau}_{X}(t) = \begin{cases} 1 & \text{for } |t| \leqslant \tau; \text{(GK)} \\ 1 - \frac{|t|}{\tau} & \text{for } |t| \leqslant \tau; \text{(HE)} \end{cases}. \tag{2.5}$$

Although in the large-time limit both the HE and the GK formula yield an unbiased estimator of the Onsager coefficient, it is important to note that the variance associated with the HE integral is a third of the GK one's [19]. Since the variance of Eq. (2.3) decreases with the total length of the trajectory, τ_{tot}, the statistical uncertainty of the GK and HE integrals, Eq. (2.4), are

$$\mathrm{var}(\widehat{L}^{X}_{ab}(\tau)) = (V\beta\epsilon)^2 \sum_{n=0}^{N} \mathrm{var}\left(\widehat{C}^{n}_{ab}\right) \Theta^{\tau}_{X}(n*\epsilon)^2 \tag{2.6}$$

$$\propto \frac{1}{\tau_{tot}} \int_{-\tau}^{\tau} \Theta^{\tau}_{X}(t)^2\,dt = \begin{cases} \frac{2\tau}{\tau_{tot}} & \text{(GK)} \\ \frac{1}{3}\frac{2\tau}{\tau_{tot}} & \text{(HE)} \end{cases}. \tag{2.7}$$

It can be proved that the prefactor in front of Eq. (2.7) is $2L^2_{ab}$ [19].

I simulate molten CsF modeled with a sample of 512 atoms interacting via the Born-Mayer-Huggins-Tosi-Fumi force field [47–49]. The trajectories are integrated via the velocity-Verlet algorithm with a time-step of 1 fs, as implemented in the `LAMMPS` code [50]. Before the microcanonical production run, I equilibrate the system in the NPT ensemble for 200 ps and in the NVT ensemble for 200 ps, at the target temperature and pressure. The Coulomb interaction is considered via Ewald summation with accuracy 10^{-7} as implemented in `LAMMPS` [50]. From a microcanonical 40 ns-long trajectory of molten CsF at 1200 K and 0 bar I compute the electric conductivity,

$$\sigma = \frac{V}{3k_B T} \int_0^\infty \left\langle \widehat{\boldsymbol{J}}(t)\cdot\widehat{\boldsymbol{J}}(0)\right\rangle_0 dt\,, \tag{2.8}$$

—$\boldsymbol{J}$ being the electric current— via GK and HE methods, and estimate the statistical uncertainty via block analysis on 1 ns-long segments. Block analysis prescribes to divide the trajectory into several segments of the same length. On each block I compute Eq. (2.4) and then evaluate the variance across the estimated samples.

In Fig. 2.1, the GK and HE integrals rapidly converge to the asymptotic value as a function of the integration time, τ. The lower panel of Fig. 2.1 shows that Eq. (2.7) grows linearly in τ, and that the variance associated with the HE formula is smaller than the GK's one. Furthermore, Eq. (2.7) demonstrates that $\mathrm{var}(\widehat{L}^{X}_{ab}(\tau))$ is proportional to $\frac{\tau}{\tau_{tot}}$. I test Eq. (2.7) numerically, computing the variance of σ from trajectories of different length. In Fig. 2.2, $\mathrm{var}\left(\sigma^{HE}(\tau)\right)\tau_{tot}$

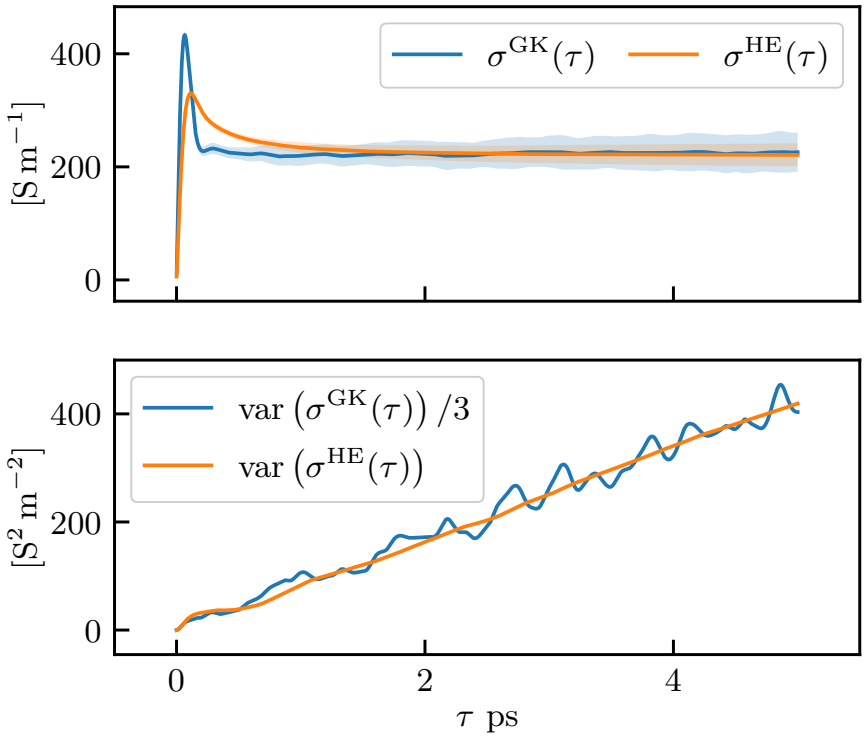

Figure 2.1: Estimate and variance of electric conductivity (upper-lower panels) of molten CsF at 1200 K and 0 bar, computed via the GK and HE integrals, Eq. (2.4) (2.7), as a function of the upper time limit, τ.

is in good agreement with the theoretical behavior, $\frac{4}{3}\sigma^2\tau$. Fig. 2.2 also displays that, when integrated over the whole trajectory, the variance of $\widehat{L}^{X}_{ab}(\tau_{tot})$ is independent on τ_{tot}. The time integration of the correlation of the fluxes, $\widehat{L}^{X}_{ab}(\tau_{tot})$, is a *non-consistent* estimator, i.e. its statistical uncertainty does not vanish in the large N-limit. At the cost of long EMD simulations, the consistency of $\widehat{L}^{X}_{ab}(\tau)$ is restored via block analysis.

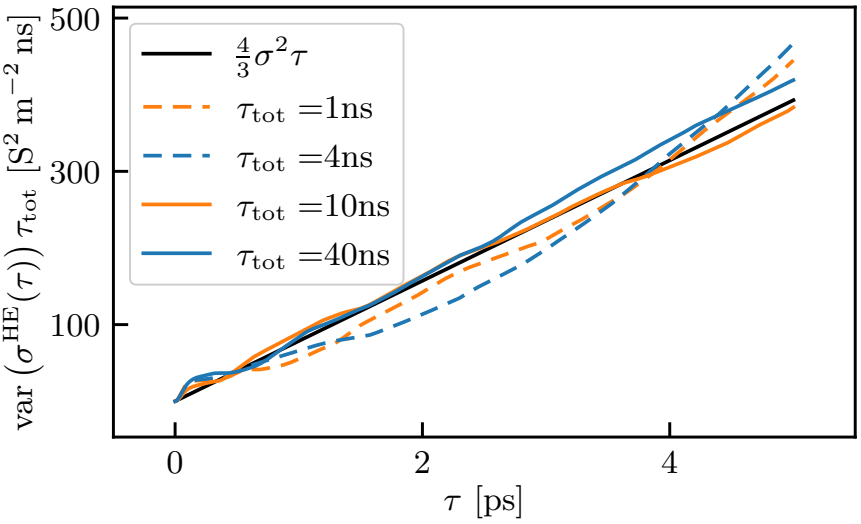

Figure 2.2: Variance of the electric conductivity of molten CsF at 1200 K and 0 bar, computed via the HE formula from trajectories of different length, τ_{tot}, compared to the theoretical behavior described in Eq. (2.7)

2.2 CEPSTRAL ANALYSIS OF DIAGONAL ONSAGER COEFFICIENTS

In the following sections, I will describe spectral methods for the computation of Onsager coefficients which do not require block analysis [15, 19, 31, 44–46]. The Power Spectral Density (PSD) is defined as the Fourier transform of the time-correlation function of a diffusive flux, J,

$$S(\omega) = \int_{-\infty}^{\infty} \left\langle \widehat{J}(t)\widehat{J}(0) \right\rangle_0 e^{i\omega t}\, dt\,. \tag{2.9}$$

The diagonal Onsager coefficient associated with the flux J, is:

$$L = \frac{V\beta}{2} \lim_{\omega\to 0} S(\omega). \tag{2.10}$$

In the 1930s, Wiener and Khinchin proved [51, 52] that the PSD of a stationary processes is the expectation value of the modulus squared of the Fourier transform of the process,

$$\begin{aligned} S(\omega) &= \lim_{\tau\to\infty} \frac{1}{\tau} \left\langle \left| \int_0^\tau \widehat{J}(t) e^{i\omega t}\, dt \right|^2 \right\rangle_0 \\ &= \lim_{\tau\to\infty} 2\,\mathrm{Re} \int_0^\tau \left\langle \widehat{J}(t)\widehat{J}(0) \right\rangle_0 e^{i\omega t}\, dt\,. \end{aligned} \tag{2.11}$$

The current time-series, $\widehat{J}_n = \widehat{J}(n\epsilon)$, $0 \leqslant n \leqslant N$, is a stationary process. Its Fourier transform,

$$\widehat{F_k} = \sum_{n=0}^{N-1} e^{2\pi i \frac{kn}{N}} \widehat{J}_n \tag{2.12}$$

is a stochastic process uncorrelated across different frequencies. Thanks to the Wiener and Khinchin theorem, Eq. (2.11), the variance of Eq.(2.12), is the estimator of the PSD, i. e. the *periodogram* [51, 52]:

$$\widehat{S}_k = S(\omega_k)\widehat{\xi}_k \tag{2.13}$$

where $k \in [0, N-1]$, $\omega_k = 2\pi\frac{k}{N}$ and $\widehat{\xi}_k$ is a χ^2-distributed random variable. Since at $k = 0$ and at the Nyquist frequency Eq. (2.12) is real, one has: $\widehat{\xi}_k \sim \frac{1}{2}\chi^2_2$ for $k \notin \{0, \frac{N}{2}\}$ and $\widehat{\xi}_k \sim \chi^2_1$ for $k \in \{0, \frac{N}{2}\}$. Notice that the periodogram, Eq. (2.13), is affected by multiplicative noise.

By performing block analysis on ℓ independent and equivalent fluxes time-series, the PSD reads

$$\begin{aligned} {}^{\ell}\widehat{S}_k &= \frac{1}{\ell}\sum_{i=1}^{\ell}\widehat{S}^i_k \\ &= S(\omega_k)\frac{1}{\ell}\sum_{i=1}^{\ell}\widehat{\xi}^i_k = S(\omega_k)\,{}^{\ell}\widehat{\xi}_k. \end{aligned} \tag{2.14}$$

Block analysis effectively reduces the error: $\operatorname{var}\left({}^{\ell}\widehat{\xi}_k\right) = \frac{1}{2\ell}$. In practice, it is a computationally expensive procedure which requires long trajectories.

2.2.1 *Cepstral analysis*

The key idea of cepstral analysis is to compute the logarithm of Eq. (2.14), i.e. the *log-periodogram*, thus turning the multiplicative noise into an additive one. Under the assumption that the log-periodogram is a smooth function of the frequency, I can apply a low-pass filter and then select only few Fourier components. I will show that the reconstructed log-periodogram yields a consistent estimator of the log-PSD at all frequencies [19, 31, 44].

The log-periodogram is

$$\begin{aligned} {}^{\ell}\widehat{\mathcal{L}}_k &= \log\left({}^{\ell}\widehat{S}_k\right) \\ &= \log\left(S(\omega_k)\right) + \log\left({}^{\ell}\widehat{\xi}_k\right). \end{aligned} \tag{2.15}$$

On one hand, Eq. (2.15) is now affected by additive noise. On the other hand, it is a biased estimator of the true log-PSD. The logarithm of a χ-squared variable can be simply decomposed in its average, λ_ℓ, and a zero-mean stochastic process,

$$\log\left({}^{\ell}\widehat{\xi}_k\right) = \lambda_\ell + {}^{\ell}\widehat{\lambda}_k, \tag{2.16}$$

where λ_ℓ and the variance of ${}^{\ell}\widehat{\lambda}_k$ can be computed analytically as

$$\lambda_\ell = \left\langle \log\left({}^{\ell}\widehat{\xi}_k\right)\right\rangle = \psi(\ell) - \log(\ell) \tag{2.17}$$

$$\sigma^2_\ell = \left\langle \log\left({}^{\ell}\widehat{\xi}_k\right)^2\right\rangle - \lambda^2_\ell = \psi'(\ell), \tag{2.18}$$

and $\psi(\ell)$, $\psi'(\ell)$ are the digamma and tri-gamma function respectively [53].

A low pass filter reduces the non-Gaussian noise of Eq. (2.15). In order to implement this procedure, I define the *cepstrum* as the inverse Fourier transform of the log-periodogram [19, 31, 44]:

$$ {}^{\ell}\widehat{C}_n = \frac{1}{N}\sum_{k=0}^{N-1} {}^{\ell}\widehat{\mathcal{L}}_k e^{-2\pi i \frac{kn}{N}}. \tag{2.19} $$

Thanks to the central limit theorem, each cepstral coefficient, Eq. (2.19), is affected by additive Gaussian-distributed noise

$$ {}^{\ell}\widehat{C}_n = \lambda_\nu \delta_{p0} + C_n + {}^{\ell}\widehat{\mu}_n \tag{2.20} $$

$$ \left\langle {}^{\ell}\widehat{\mu}_n{}^2 \right\rangle = \begin{cases} \frac{\sigma_\ell^2}{N} & n \notin \{0, \frac{N}{2}\} \\ \frac{2\sigma_\ell^2}{N} & n \in \{0, \frac{N}{2}\} \end{cases}. \tag{2.21} $$

In the large-N limit, ${}^{\ell}\widehat{\mu}_n$s are independent zero-mean normal variables. Cepstral coefficients are real, periodic and symmetric: ${}^{\ell}\widehat{C}_n = {}^{\ell}\widehat{C}_{N-n}$. Provided that the log-PSD is smooth enough, only P cepstral coefficients are required to reconstruct the original spectrum: ${}^{\ell}\widehat{C}_n \approx 0$ for $P \leqslant n \leqslant N - P$. The selection of the meaningful cepstral coefficient is a crucial step in the cepstral procedure. If one selects blindly all ${}^{\ell}\widehat{C}_n$s, then the estimator would be non-consistent for the same reasons Eq. (2.4) is not. I am interested in the $k = 0$ component of Eq. (2.15) that reads

$$ {}^{\ell}\widehat{\mathcal{L}}_0(P) = {}^{\ell}\widehat{C}_0 + 2\sum_{n=1}^{P-1} {}^{\ell}\widehat{C}_n \tag{2.22} $$

$$ = \lambda_\ell + \log(S(0)) + {}^{\ell}\widehat{\mu}_0 + 2\sum_{n=1}^{P-1} {}^{\ell}\widehat{\mu}_n. \tag{2.23} $$

Therefore, the expectation value and variance of Eq. (2.22) are

$$ {}^{\ell}\mathcal{L}_0(P) = \left\langle {}^{\ell}\widehat{\mathcal{L}}_0(P) \right\rangle \qquad = \lambda_\ell + \log(S(0)) \tag{2.24} $$

$$ \sigma_\ell(P,N)^2 = \operatorname{var}\left({}^{\ell}\widehat{\mathcal{L}_0}(P) \right) = \sigma_\ell^2 \frac{4P-2}{N}. \tag{2.25} $$

There are several criteria for selecting the optimal number of cepstral coefficients. A common choice is the Akaike Information Criterion (AIC), which aims to balance the bias introduced by the finite summation in Eq. (2.22) and the variance in Eq. (2.25). AIC selects the number of cepstral coefficients which minimizes the loss function,

$$ \mathrm{AIC}(P) = -2\max_{\mathbf{w}} \log \mathfrak{L}(\mathbf{w}, P) + 2P, \tag{2.26} $$

where $\mathfrak{L}(\mathbf{w}, P)$, $\mathbf{w} = \{w_0 \ldots w_{P-1}\}$, is the Likelihood function:

$$ 2\log \mathfrak{L}(\mathbf{w}, P) = -(w_0 + \lambda_\ell - {}^{\ell}\widehat{C}_0)^2 - \sum_{n=1}^{P-1}(w_n - {}^{\ell}\widehat{C}_n)^2 - \sum_{n=P}^{N/2}({}^{\ell}\widehat{C}_n). \tag{2.27} $$

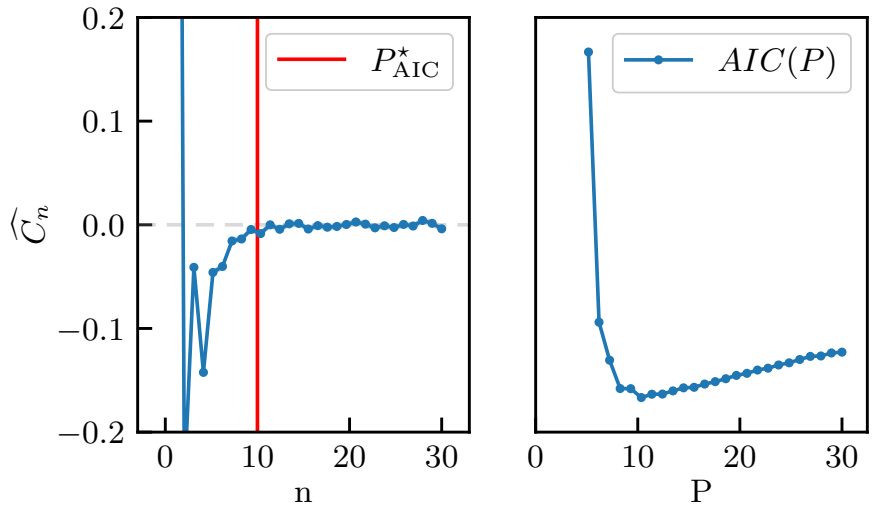

Figure 2.3: Left panel: Cepstral coefficients of the electric current periodogram obtained from a 1 ns-long microcanonical trajectory of molten CsF at 1200 K and 0 bar. The red vertical line highlights the best number of cepstral coefficients selected by the AIC. Right panel: AIC function computed as a function of the number of cepstral coefficients on the same trajectory

The best number of parameters, $P^{\star}$ is estimated as

$$P^{\star} = \underset{P}{\operatorname{argmin}}\ \mathrm{AIC}(P). \tag{2.28}$$

Eq. (2.27) is maximized when $w_n = {}^{\ell}\widehat{C}_n - \lambda_{\ell}\delta_{n0}$, thus the AIC function reads

$$\mathrm{AIC}(P) = \sum_{n=P}^{N/2} ({}^{\ell}\widehat{C}_n)^2 + 2P. \tag{2.29}$$

Cepstral analysis combined with AIC yields a consistent estimator of diagonal Onsager coefficients [31, 44, 54]. Therefore, cepstral analysis does not require block analysis to estimate the variance of the filtered spectrum [31, 44, 54]. By avoiding block analysis, the numerical efficiency of computing transport coefficients improves [28, 31, 32, 44, 54–57].

2.2.2 *Numerical experiments on cepstral analysis*

Cepstral analysis can be employed to compute electric conductivity, Eq. (2.8), from a 1 ns trajectory of molten CsF at 1200 K and 0 bar. The results are displayed in Fig. 2.3 and Fig. 2.4. In Fig. 2.3 cepstral coefficients rapidly converge to zero, and AIC selects only 10 of them. As shown in Fig. 2.4, the variance of Eq. (2.24) grows linearly with P. The cepstral estimator of diagonal Onsager coefficients, Eq. (2.24) and Eq. (2.25), is itself a Gaussian random variable. In order to showcase this statement, I evaluate electric conductivity of CsF via

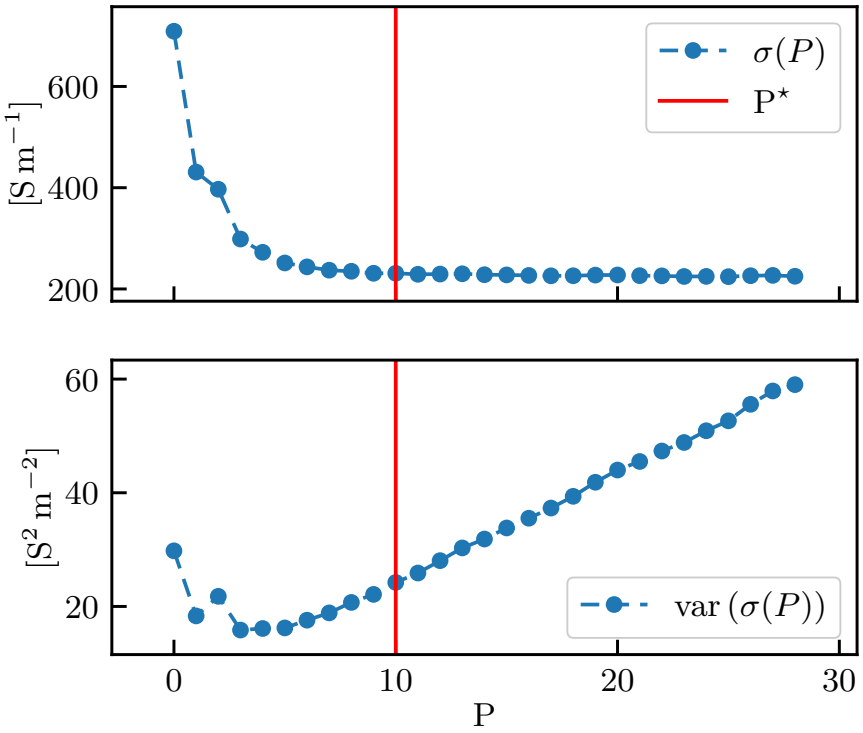

Figure 2.4: Cepstral estimate of the electric conductivity (upper panel) and its variance (lower panel), computed from a 1 ns-long microcanonical trajectory of molten CsF at 1200 K and 0 bar, as a function of the number of cepstral coefficients considered.

cepstral analysis from 200 200 ps-long microcanonical trajectories and then compute the histogram of the outcomes. The results are displayed in Fig. 2.5 and compared to the reference cepstral estimate obtained from a 1 ns-long trajectory. The red curve is a Gaussian fit to the data and the green one is also Gaussian, with mean and standard deviation evaluated averaging those predicted by cepstral analysis over the 200 segments. The data in Fig. 2.5 also pass the Shapiro-Wilk normality test at a standard level of significance of 0.05 [57, 58].

Thermal conductivity of multicomponent systems via cepstral analysis

Cepstral analysis is a viable tool to compute thermal conductivity in both single and multicomponent systems. In multicomponent systems thermal conductivity is computed as Eq. (1.76). The PSD-matrix, $\mathbf{S}(\omega)$, is

$$\begin{aligned} S^{ij}(\omega) &= \lim_{\tau\to\infty} \frac{1}{\tau} \operatorname{Re} \left\langle \left(\int_0^\tau \widehat{J}_i(t) e^{i\omega t}\, dt \right) \left(\int_0^\tau \widehat{J}_j(t) e^{-i\omega t}\, dt \right) \right\rangle_0 \\ &= \lim_{\tau\to\infty} 2 \operatorname{Re} \int_0^\tau \left\langle \widehat{J}_i(t) \widehat{J}_j(0) \right\rangle_0 e^{i\omega t}\, dt\,, \end{aligned} \tag{2.30}$$

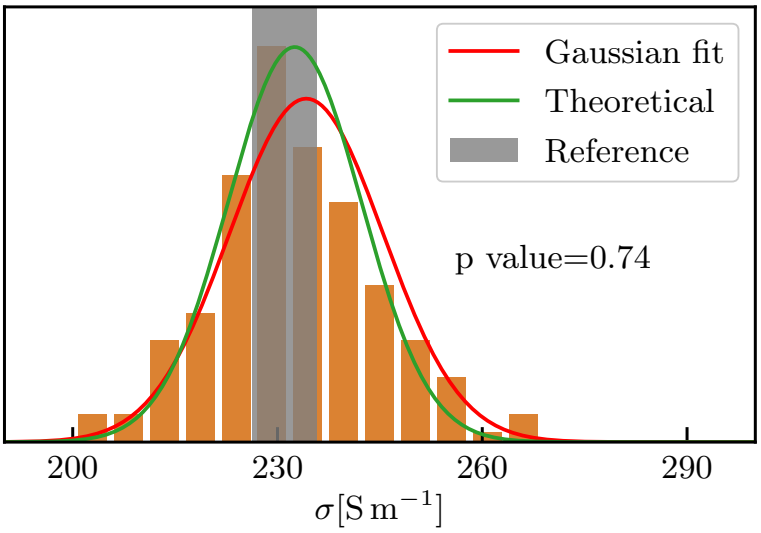

Figure 2.5: Histogram of the cepstral estimates of electric conductivity of molten CsF at 1200 K and 0 bar, computed from 200 200 ps-long trajectories. The dashed grey area is the electric conductivity and its statistical uncertainty estimated via cepstral analysis from a 1 ns-long simulation. The red curve is a Gaussian fit on the data and the green one is a Gaussian whose mean and variance are the average of the cepstral ones estimated from the 200 trajectories. In the plot is also reported the p-value of the Shapiro-Wilk normality test.

where $i, j \in \{E, 1 \dots M\}$ and M is the number of diffusive fluxes [31, 46]. The expression for thermal conductivity in multicomponent systems is

$$\kappa = \frac{V}{2k_B T^2} \lim_{\omega \to 0} \bar{S}^E(\omega) = \frac{V}{2k_B T^2} \lim_{\omega \to 0} \frac{1}{(\mathbf{S}^{-1}(\omega))_{EE}}, \tag{2.31}$$

where $\bar{S}^E(\omega)$ is the reciprocal of the EE-element of the inverse periodogram matrix, Eq.(2.30). $\bar{S}^E(\omega)$ is referred to as Shur complement, as defined in Eq. (1.76) [31]. Eq. (2.31) derives directly from the thermodynamic definition of thermal conductivity:

$$\kappa \equiv \left. \frac{\left\langle \widehat{J}_e \right\rangle}{\nabla T} \right|_{J_k = 0}.$$

In order to apply cepstral analysis, Eq. (2.31) must be a positive definite χ^2-distributed variable.

The periodogram-matrix $\widehat{\mathbf{S}}_k$, the estimator of Eq. (2.30), is a Wishart process [31]. Interestingly, the Shur complement of a Wishart matrix is itself Wishart distributed [31]. Therefore, the estimator of Eq. (2.31) is χ^2-distributed and the cepstral filter can be applied [31]. In order to compute the correct transport coefficient in liquids with more than one diffusing species, the multicomponent formula for thermal conductivity is crucial. To showcase this statement,

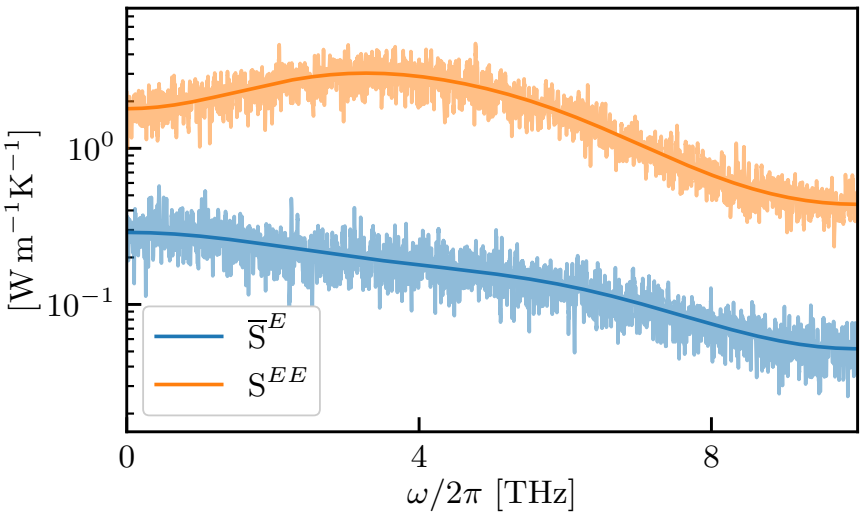

Figure 2.6: Cepstral filtered PSD (solid line) and raw periodogram (shaded area) of the energy flux (blue) and of Eq. (2.31) (orange), computed from a 1 ns-long trajectory of molten CsF at 1200 K and 0 bar.

I compute thermal conductivity of molten CsF from a 1 ns-long microcanonical trajectory, according to Eq. (2.31) and to the single-component formula,

$$\kappa = \frac{V}{2k_B T^2} \lim_{\omega \to 0} S^{EE}(\omega). \tag{2.32}$$

The outcomes of the analysis are shown in Fig. 2.6 and prove numerically that the single-component formula significantly deviates from Eq. (2.31). In the present test case, Eq. (2.32) overestimates thermal conductivity by about one order of magnitude.

The cepstral filter is based on the statistical properties of the diagonal elements of the periodogram matrix. AIC automatically selects the optimal number of cepstral coefficients in Eq. (2.24) and Eq. (2.25), yielding a consistent estimator of the Onsager coefficient. Bypassing block analysis, cepstral analysis requires short EMD simulations to achieve reasonable error bars [31, 44]. However, cepstral analysis has two main drawbacks: it is only applicable to χ^2-squared random variables and it performs poorly in filtering zero-frequency peaked PSD [44], due to the large number of cepstral coefficients selected by AIC [44, 46]. For example, in weakly anharmonic crystals, because of the AIC, cepstral estimates significantly deviate from the GK ones. The Minimum Mean Square Error (MMSE) model selection method was developed as an alternative to the AIC one, and partially alleviates the numerical drawbacks of cepstral analysis in such systems [59]. Remarkably, the cepstral filter is not applicable to off-diagonal Onsager coefficients.

2.3 BAYESIAN REGRESSION ANALYSIS OF WISHART PROCESSES

Cepstral analysis is a powerful tool for estimating diagonal Onsager coefficients, such as electric conductivity [54], shear and bulk viscosity [57] and thermal conductivity [28, 31, 32, 54–56]. Nevertheless, this filtering technique is not designed for off-diagonal Onsager coefficients. The periodogram matrix, $\widehat{\mathbf{S}}_k$, —i. e. the estimator of Eq. (2.30)—

$$\begin{aligned}\mathbf{S}(\omega_k) &= \left\langle \widehat{\mathbf{S}}_k \right\rangle, \qquad (2.33)\\ S^{ij}(\omega_k) &= \lim_{\tau\to\infty} \frac{1}{\tau} \operatorname{Re}\left\langle \left(\int_0^\tau \widehat{J}_i(t) e^{i\omega_k t}\, dt\right)\left(\int_0^\tau \widehat{J}_j(t) e^{-i\omega_k t}\, dt\right)\right\rangle_0 \\ &= \lim_{\tau\to\infty} 2\operatorname{Re}\int_0^\tau \left\langle \widehat{J}_i(t)\widehat{J}_j(0)\right\rangle_0 e^{i\omega_k t}\, dt,\end{aligned}$$

is a Wishart-distributed stochastic process. The Onsager matrix, $\mathbf{L}$, is the 0-frequency limit of the periodogram matrix:

$$\mathbf{L} = \frac{V\beta}{2} \lim_{\omega\to 0} \mathbf{S}(\omega).$$

$\widehat{S}_k^{ij}$s are neither positive definite random variables nor χ^2-squared distributed. For these reasons one cannot apply cepstral analysis to filter its noise.

We devise a tailored filtering method to address this type of stochastic processes [45, 46]. The off-diagonal periodogram, $\widehat{S}_k^{ij}$, is a stochastic process which follows the Gamma-Variance distribution, whose probability density function is [60]

$$\begin{aligned}p\left(S_k^{ij}\right) =& \frac{\left|S_k^{ij}\right|^{\frac{\nu-1}{2}}}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{2^{\nu-1}\pi(1-\rho(\omega_k)^2)}[S^{ii}(\omega_k)S^{jj}(\omega_k)]^{\frac{\nu+1}{4}}} \\ &\times \mathcal{K}_{\frac{\nu-1}{2}}\left(\frac{\left|S_k^{ij}\right|}{[S^{ii}(\omega_k)S^{ii}(\omega_k)]^{1/2}(1-\rho(\omega_k)^2)}\right) \qquad (2.34)\\ &\times \exp\left\{\frac{\rho(\omega_k)S_k^{ij}}{[S^{ii}(\omega_k)S^{jj}(\omega_k)]^{1/2}(1-\rho(\omega_k)^2)}\right\},\end{aligned}$$

where $\nu = 2\ell$, ℓ is the number of equivalent fluxes (in isotropic fluids $\ell = 3$), $\rho(\omega_k)$ is the correlation coefficient, Γ is the Gamma function and $\mathcal{K}$ is the modified Bessel function of second kind. After a simple change of variables, the correlation coefficient estimator, $\widehat{\rho_k}$, and its probability density are

$$\widehat{\rho}_k = \frac{\widehat{S}_k^{ij}}{[S^{ii}(\omega_k)S^{jj}(\omega_k)]^{1/2}} \qquad (2.35)$$

$$p(\rho_k) \propto \frac{|\rho_k|^{\frac{\nu-1}{2}}}{\sqrt{1-\rho(\omega_k)^2}} \mathcal{K}_{\frac{\nu-1}{2}}\left(\frac{|\rho_k|}{1-\rho(\omega_k)^2}\right) \exp\left\{\frac{\rho(\omega_k)\rho_k}{1-\rho(\omega_k)^2}\right\}. \qquad (2.36)$$

Eq. (2.35) describes a Gamma-Variance stochastic process, independent at different frequencies [51, 52]. In order to obtain a reliable and accurate estimate of the PSD, we employ Bayesian regression [45, 46].

2.3.1 *Bayesian regression analysis*

Let us consider two stochastic processes $\widehat{X}$ and $\widehat{Y}$. Under the hypothesis $p(X,Y) = p(Y,X)$, the Bayes theorem states

$$p(Y)\,p(X \mid Y) = p(X)\,p(Y \mid X). \tag{2.37}$$

The aim of this section is to describe how to apply Bayesian regression methods to filter the periodogram.

Given a dataset, $\mathcal{D}$, and a model, $\Theta(\cdot,\cdot)$, defined by P parameters $\mathbf{w} \equiv \{w_0 \dots w_{P-1}\}$, Bayesian regression searches for the best model that approximates the dataset, $\mathcal{D}$. $\mathcal{D} \equiv \{y_n; x_n\}$ is a collection of samples of a function $y(x)$ which depends on an input x. The dataset itself is a stochastic process that is distributed according to $p(\mathcal{D} \mid y)$,

$$\langle y_n \rangle_{p(\mathcal{D}|y)} = y(x_n). \tag{2.38}$$

Here, $\langle\cdot\rangle_{p(\circ)}$ represents the expectation value with respect to the probability distribution $p(\circ)$. Of course, $y(x)$ is unknown. The aim of the procedure is to seek the optimal set of parameters, $\overline{\mathbf{w}}$, such that $\Theta(\overline{\mathbf{w}}; x) \approx y(x)$. The first step in the Bayesian procedure for determining the best model, is to define the probability that the dataset was generated not by $y(x)$, but by $\Theta(\mathbf{w},\cdot)$: the so-called *Likelihood distribution function*, $p(\mathcal{D} \mid \mathbf{w})$,

$$\langle y_n \rangle_{p(\mathcal{D}|\mathbf{w})} = \Theta(\mathbf{w}; x_n). \tag{2.39}$$

The Likelihood function characterizes how the dataset is distributed knowing the model. Bayes theorem reverses the flow of information, thus allowing to compute the distribution of the parameters at given dataset:

$$p(\mathbf{w} \mid \mathcal{D}) = \frac{p(\mathcal{D} \mid \mathbf{w})\,p(\mathbf{w})}{p(\mathcal{D})}, \tag{2.40}$$

where $p(\mathbf{w})$ is the *prior* distribution of the parameters, i.e. the agnostic probability density of $\mathbf{w}$. $p(\mathbf{w} \mid \mathcal{D})$ is usually referred to as *posterior* distribution of the parameters. The optimal value of the parameters, provided the information stored in the dataset, is estimated as the expectation value of $\mathbf{w}$ over the posterior distribution [61]:

$$\langle \mathbf{w} \rangle_{p(\mathbf{w}|\mathcal{D})} = \overline{\mathbf{w}}. \tag{2.41}$$

There is significant flexibility in selecting the model, which can be as expressive as needed, as well as in choosing the prior. For instance, the prior can be a Gaussian parametrized by a positive real number α:

$$p(\mathbf{w} \mid \alpha) = \left(\frac{\alpha}{2\pi}\right)^{\frac{P}{2}} \exp\left\{-\frac{\alpha}{2}\|\mathbf{w}\|^2\right\}. \tag{2.42}$$

This prescription is equivalent to a $\mathcal{L}^2$ regularization term [61]. Bayesian regression allows to select the optimal prior by computing its posterior distribution, $p(\alpha \mid \mathcal{D})$. According to Bayes theorem, one has [61]

$$p(\alpha \mid \mathcal{D}) \propto p(\mathcal{D} \mid \alpha)\, p(\alpha) \tag{2.43}$$

$$p(\mathcal{D} \mid \alpha) = \int p(\mathcal{D} \mid \mathbf{w})\, p(\mathbf{w} \mid \alpha)\, d\mathbf{w}\,. \tag{2.44}$$

If the log-Likelihood function is quadratic, the model is linear in the parameters and, the prior is Gaussian, Eq. (2.44) can be computed analytically [61]. On the contrary, if the Likelihood is not Gaussian, only numerical solutions are available. As in Eq. (2.44) and Eq. (2.43), the best number of parameters of the model, P, is estimated maximizing

$$p(P \mid \mathcal{D}) \propto p(\mathcal{D} \mid P)\, p(P) \tag{2.45}$$

$$p(\mathcal{D} \mid P) = \int p(\mathcal{D} \mid \mathbf{w}, P)\, p(\mathbf{w} \mid P)\, d\mathbf{w}\,. \tag{2.46}$$

In the large sample limit, the integrand in Eq. (2.46), $p(\mathcal{D} \mid \mathbf{w}, P)\, p(\mathbf{w} \mid P) \equiv \widetilde{p}(\mathbf{w} \mid \mathcal{D}, P)$, can be rewritten as [61]

$$\begin{aligned} \log \widetilde{p}(\mathbf{w} \mid \mathcal{D}) &\approx \log \widetilde{p}(\overline{\mathbf{w}} \mid \mathcal{D}) + (\mathbf{w} - \overline{\mathbf{w}})^{\dagger}\, \frac{\partial^2 \log \widetilde{p}(\mathbf{w} \mid \mathcal{D})}{2\partial \mathbf{w}^2}\bigg|_{\overline{\mathbf{w}}} (\mathbf{w} - \overline{\mathbf{w}}) \\ &= \log \widetilde{p}(\overline{\mathbf{w}} \mid \mathcal{D}) - \frac{N}{2}\, (\mathbf{w} - \overline{\mathbf{w}})^{\dagger}\, \Im(\overline{\mathbf{w}})\, (\mathbf{w} - \overline{\mathbf{w}})\,, \end{aligned} \tag{2.47}$$

where $\Im(\overline{\mathbf{w}})$ is the average Fisher information matrix over the dataset [61]. As a general remark I want to stress that, assuming reasonable hypotheses on the prior distribution of the parameters, i.e. smoothness etc., the Bernstein-Von Mises theorem states that in the large-sample limit, the posterior distribution converges to a Gaussian, centered in $\overline{\mathbf{w}}$, with covariance matrix $N^{-1}\Im^{-1}(\overline{\mathbf{w}})$ [62]. This simple statement proves the consistency of the Bayesian regression approach. Integrating Eq. (2.47) and assuming a uniform prior of the models, $p(P)$, I recover the Bayesian Information Criterion (BIC) [61]:

$$\log p(P \mid \mathcal{D}) \approx \log p(\mathcal{D} \mid \overline{\mathbf{w}}) - \frac{k}{2} \log N \equiv -2\mathrm{BIC}(P). \tag{2.48}$$

2.3.2 *Numerical experiments on Bayesian regression analysis*

The aforesaid Bayesian method has been applied to two "synthetic" datasets, each consisting of 8000 samples of Gamma-Variance distributed stochastic processes, $y(x_k)$. The synthetic datasets were generated from simple functions, sin and tanh, with Gamma-variance noise added to create spectra compatible with

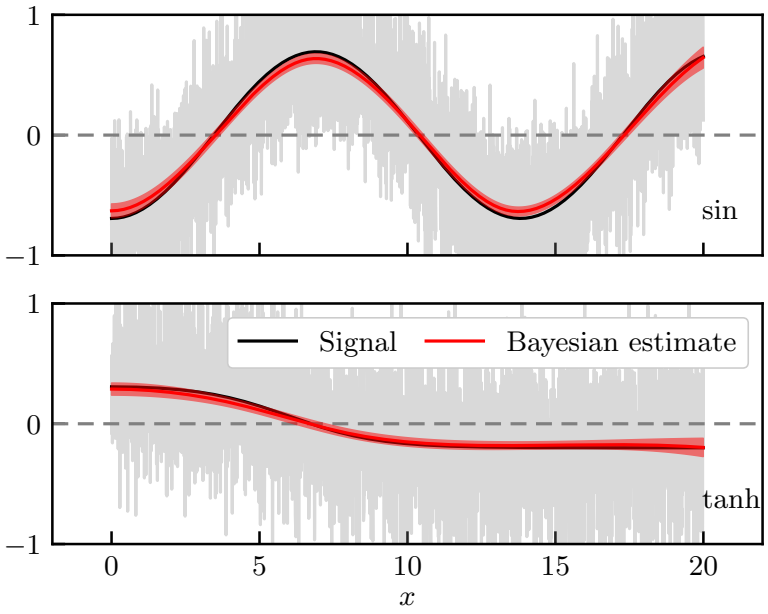

Figure 2.7: The grey shaded area is the synthetic dataset, the solid black line displays the signal. The solid red line and the red shaded area show the Bayesian regression and its estimated statistical uncertainty. The upper and lower panels address the synthetic datasets generated via sin and tanh functions, respectively.

bona fide off-diagonal periodograms. The datasets are modeled with a cubic spline [63], Θ_{cs},

$$y(x_k) \approx \Theta_{cs}(x_k|\mathbf{w}) = \begin{cases} C_1(x_k), \; x^0 = 0 \leqslant x_k \leqslant x^1 \\ \vdots \\ C_i(x_k), \; x^{i-1} \leqslant x_k \leqslant x^i \\ \vdots \\ C_P(x_k), \; x^{P-1} \leqslant x_k \leqslant x^M \end{cases} \tag{2.49}$$

where $C_i(x)$ are third order polynomials, with $C_i(x^{i-1}) \doteq w_{i-1}$; $\Theta_{cs}(x_k|\mathbf{w})$ is twice differentiable with respect to its argument, x_k; $\mathbf{w} = \{w_0, \ldots, w_{P-1}\}$ represent the P parameters of the model optimized through Bayesian regression; the inputs $x^0, \ldots, x^{P-1}$ are the (fixed) values at which the spline is evaluated, usually called *knots*. Notice the superscript labels that distinguish the knots from the sample inputs.

The Likelihood function, that is the distribution of the data knowing the model, is defined in Eq. (2.36). The optimal parameters are estimated via Eq. (2.41), computing the expectation value as a Monte Carlo Markov Chain (MCMC) average over the posterior distribution of the parameters, assuming a uniform prior, as implemented in the `emcee` code [64, 65]. The results of the Bayesian regression are displayed in Fig. 2.7, which shows that the spline

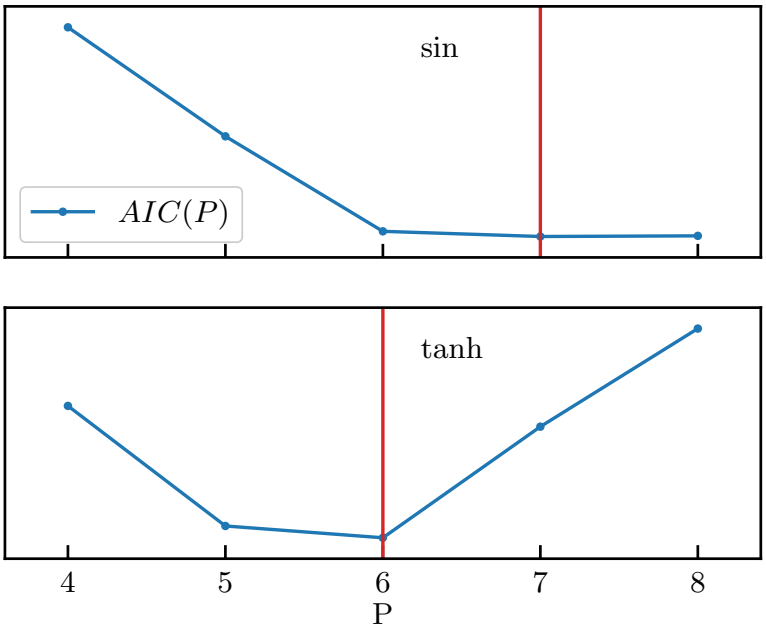

Figure 2.8: Eq. (2.26) as a function of the number of the cubic spline knots. The vertical red line highlights the minimum of the AIC function and indicated the best model selected by the AIC. The upper and lower panels address the synthetic datasets generated via sin and tanh functions, respectively.

model is expressive enough to capture the features of the signal. In order to select the best number of parameters I apply AIC, Eq (2.26). The AIC function, Eq. (2.26), is displayed in Fig. 2.8 for both datasets. AIC always selects a low number of parameters, proving that a higher number of parameters would only increase the variance on the prediction. The parameters, drawn from the MCMC, pass the Shapiro-Wilk test at a level of significance of 0.05. The impact of a Gaussian prior on the results has been studied. The optimal value of α is selected by maximizing Eq. (2.44), under the assumption of a uniform $p(\alpha)$, as introduced in Eq. (2.43). The right panels of Fig. 2.9 display $\log p(\alpha \mid \mathcal{D})$ as a function of the width of the Gaussian prior and its associated maximum. The selected Gaussian priors do not alter the outcomes obtained with a uniform one.

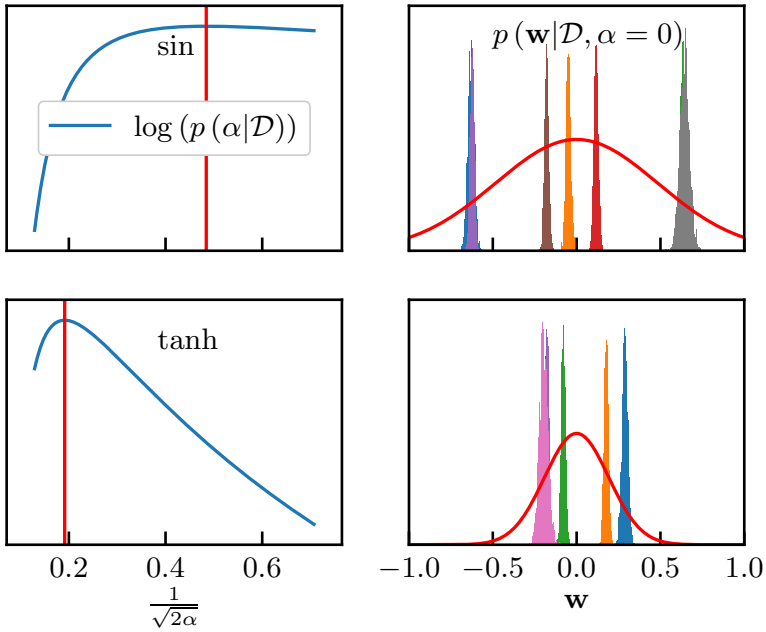

Figure 2.9: Left panels: $\log p(\alpha \mid \mathcal{D})$ estimated via Eq. (2.44) assuming a uniform distribution $p(\alpha)$. The vertical line indicates the maximum of $\log p(\alpha \mid \mathcal{D})$ and thus the optimal value of α. Right panels: distribution of the parameters sampled from the MCMC according to the posterior distribution, $p(\mathbf{w} \mid \mathcal{D}, \alpha = 0)$. The red solid curve is the optimal prior, $p(\mathbf{w} \mid \alpha)$, selected maximizing $\log p(\alpha \mid \mathcal{D})$. The upper and lower panels address the synthetic datasets generated via sin and tanh functions, respectively.

2.4 BAYESIAN EXTRAPOLATION OF STATIC RESPONSE FUNCTIONS

In condensed matter systems, because of the enforced PBC, only external periodic perturbations are allowed. Thus, the response to a periodic perturbation is of particular interest. In Eq. (1.82) I have showed that, in the long-wavelength limit, some static response functions converge to well-defined thermodynamic derivatives.

In order to extrapolate the long wavelength limit, I design a Bayesian linear regression method. Let us consider the spatial Fourier transform, as in $f(\mathbf{k}) = \frac{1}{V}\int_V f(\mathbf{r})e^{-i\mathbf{k}\cdot\mathbf{r}}\,d\mathbf{r}$, of the density of a conserved quantity, $a(\mathbf{k})$ and $b(\mathbf{k})$, and a perturbation acting on the Hamiltonian: $\mathcal{H}'(\Gamma) = V v_a(\mathbf{k}) a(-\mathbf{k}, \Gamma)$. The external potential, $v_a(\mathbf{k})$, acts as a tunable parameter. At first order in $v_a(\mathbf{k})$, one has

$$\begin{aligned}\langle b(\mathbf{k})\rangle &= \frac{\int b(\mathbf{k},\Gamma)e^{-\beta(\mathcal{H}+\mathcal{H}')}\,d\Gamma}{\int e^{-\beta(\mathcal{H}+\mathcal{H}')}\,d\Gamma} \approx \frac{\int b(\mathbf{k},\Gamma)e^{-\beta\mathcal{H}}(1-\beta V v_a(\mathbf{k})a(-\mathbf{k},\Gamma))\,d\Gamma}{\int e^{-\beta\mathcal{H}}(1-\beta V v_a(\mathbf{k})a(-\mathbf{k},\Gamma))\,d\Gamma} \\ &\approx \langle b(\mathbf{k})\rangle_0 - \beta V\left(\langle b(\mathbf{k})a(-\mathbf{k})\rangle_0 - \langle b(\mathbf{k})\rangle_0\langle a(-\mathbf{k})\rangle_0\right) v_a(\mathbf{k}).\end{aligned} \tag{2.50}$$

The static linear response function $\chi_{ab}(\mathbf{k})$ is

$$\begin{aligned}\chi_{ab}(\mathbf{k}) &\equiv \left.\frac{\partial b(\mathbf{k})}{\partial v_a(\mathbf{k})}\right|_{v_a(\mathbf{k})=0} \\ &= -\beta V\left(\langle b(\mathbf{k})a(-\mathbf{k})\rangle_0 - \langle b(\mathbf{k})\rangle_0\langle a(-\mathbf{k})\rangle_0\right).\end{aligned} \tag{2.51}$$

Static linear response theory determines the response of the thermodynamic derivatives such as the dielectric constant, specific heats and the isothermal compressibility. Because of the central limit theory, the estimator of Eq. (2.51) is a Gaussian process. The static response function, $\chi_{ab}(\mathbf{k})$ can be approximated by a low-order polynomial of the wave-vector, $\mathbf{k}$. In a liquid, isotropy dictates that $\chi_{ab}(\mathbf{k})$ is only a function of the modulus squared of its argument, k^2:

$$\chi_{ab}(\mathbf{k}) = w_0 + w_1 k^2 + \cdots + w_{P-1}k^{2P-2}, \tag{2.52}$$

where $\mathbf{w} = \{w_0, w_1, \ldots, w_{P-1}\}$ are the parameters to be optimized in the Bayesian regression.

Because of PBC, from a single EMD trajectory with a cubic simulation cell, one has access to an equispaced grid of $\mathbf{k}$-vectors in reciprocal space. The bigger the cubic simulation cell, the denser the grid becomes. Even in isotropic fluids, a finite cubic simulation cell with PBC, reduces spherical symmetry to cubic. In reciprocal space, the space Fourier transforms of the response functions are symmetric under cubic transformations. The symmetry reduction, from spherical to cubic, might affect finite size samples of wave-vector dependent response functions, making estimates at wave-vectors with the same magnitude but not equivalent under cubic symmetry, incompatible. For example, from a EMD trajectory with a cubic simulation cell of linear dimension L, finite size effects might appear at $\mathbf{k}$-vectors: $\frac{2\pi}{L}(2,2,1)$ and $\frac{2\pi}{L}(3,0,0)$. In isotropic systems, the infinite size limit of the wave-vector dependent response functions must only depend on the modulus of the probed wave-vector.

A simple Bayesian linear regression procedure can be applied to evaluate the long-wavelength limit of Eq. (2.52). Since the estimator of Eq. (2.51) is Gaussian, the Likelihood distribution function reads

$$p(\mathcal{D} \mid \mathbf{w}) = \exp\left\{-\sum_{i=1}^{N}\frac{\left(\chi^i_{ab} - \mathbf{w}\boldsymbol{\Theta}_i\right)^2}{2\sigma_i^2}\right\}, \tag{2.53}$$

where $\chi^i_{ab} = \chi_{ab}(\mathbf{k}_i)$ is the estimated response function at wave-vector $\mathbf{k}_i$, $\boldsymbol{\Theta}$ is the basis set of monomials of even degree, $\boldsymbol{\Theta}_i \equiv \boldsymbol{\Theta}(\mathbf{k}_i) = \{1, k_i^2, \ldots, k_i^{2P-2}\}$, σ_i^2 is the variance associated to χ^i_{ab} computed via block analysis. As discussed in Sec. 2.3, the posterior distribution is computed thanks to Bayes theorem, Eq. (2.40). The prior distribution of the parameters is Gaussian parametrized, by a real positive number α, Eq. (2.42). Since both the prior and the Likelihood distributions are Gaussian, the posterior is Normal with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$:

$$p(\mathbf{w} \mid \mathcal{D}, \alpha) = \left(\frac{1}{\sqrt{2\pi|\boldsymbol{\Sigma}|}}\right)^P e^{-\frac{1}{2}(\mathbf{w}-\boldsymbol{\mu})^\dagger \boldsymbol{\Sigma}^{-1}(\mathbf{w}-\boldsymbol{\mu})}. \tag{2.54}$$

After some algebra, one has [61]

$$\boldsymbol{\mu} = \boldsymbol{\Sigma}\boldsymbol{\Phi}^{\dagger}\boldsymbol{y} \tag{2.55}$$

$$\boldsymbol{\Sigma}^{-1} = \alpha\mathbb{1} + \boldsymbol{\Phi}^{\dagger}\boldsymbol{\Phi}, \tag{2.56}$$

where I define $\boldsymbol{y} \equiv \{\frac{\chi^{i}_{ab}}{\sigma_i}\}$ to be the vector of the elements in the dataset $\mathcal{D}$ and $\boldsymbol{\Phi}_i \equiv \frac{\boldsymbol{\Theta}_i}{\sigma_i}$. Furthermore, the posterior distribution of the hyper-parameter α, Eq. (2.44), is [61]

$$p(\mathcal{D} \mid \alpha) = \left(\frac{1}{2\pi}\right)^{N/2} \left(\frac{\alpha}{2\pi}\right)^{P/2} \int e^{-E(\mathbf{w})}\, d\mathbf{w} \tag{2.57}$$

$$E(\mathbf{w}) = \sum_{i=1}^{N} \frac{(y_i - \mathbf{w}\boldsymbol{\Phi}_i)^2}{2} + \frac{\alpha}{2}\|\mathbf{w}\|^2. \tag{2.58}$$

The integral in Eq. (2.57) reads

$$\left(\frac{\alpha}{2\pi}\right)^{P/2} \int e^{-E(\mathbf{w})}\, d\mathbf{w} = e^{-E(\boldsymbol{\mu})}\, (2\pi)^{P/2}\, |\boldsymbol{\Sigma}|^{1/2} \tag{2.59}$$

and the log-posterior of α is

$$\log p(\mathcal{D} \mid \alpha) = \frac{P}{2}\log\alpha - E(\boldsymbol{\mu}) - \frac{1}{2}\log\left|\boldsymbol{\Sigma}^{-1}\right| - \frac{N}{2}\log 2\pi. \tag{2.60}$$

I select the optimal prior maximizing Eq. (2.60) [61]. Under the hypothesis that $p(\alpha \mid \mathcal{D})$ is extremely peaked around its maximum, $\overline{\alpha}$, I can write

$$p(\mathcal{D} \mid P) = \int p(\mathcal{D} \mid \mathbf{w})\, p(\mathbf{w} \mid \alpha, P)\, p(\alpha)\, d\mathbf{w}\, d\alpha \tag{2.61}$$

$$= \int p(\mathcal{D} \mid \alpha, P)\, p(\alpha)\, d\alpha \tag{2.62}$$

$$\approx p(\mathcal{D} \mid \overline{\alpha}, P)\, p(\overline{\alpha}). \tag{2.63}$$

Since $p(\overline{\alpha})$ does not depend on P, the optimal number of parameters is the one which maximizes $p(\mathcal{D} \mid \overline{\alpha}, P)$.

In order to validate the long-wavelength Bayesian extrapolation of static response functions, I compute the dielectric constant, the specific heat at constant pressure and isothermal compressibility of 544 rigid `TIP4P/2005` water molecules at ambient conditions [66]. The `TIP4P/2005` force field is well known to approximate accurately some thermodynamic properties of water [66–73]. Furthermore, the `TIP4P/2005` model associates fixed point charges to the atoms. I generate 1 ns-long EMD trajectories, integrated using the velocity-Verlet algorithm with a time step of 0.5 fs as implemented in the `LAMMPS` code [50]. In the simulations, I equilibrate the system in the NPT ensemble at 350 K and 0 bar for 100 ps. After the equilibration run, from the 1 ns-long NVE trajectory, I sample heat, charge and, number density every 50 fs. The

long-wavelength limit of static response functions coincide with standard fluctuation methods [14, 74]:

$$1 - \frac{1}{\epsilon} = \frac{\partial P}{\partial D} = \frac{4\pi V}{k_B T} \lim_{\mathbf{k}\to 0} \frac{\langle \widehat{\rho}(\mathbf{k})\widehat{\rho}(-\mathbf{k})\rangle_{NVE,D}}{k^2} = \lim_{\mathbf{k}\to 0} \chi(\mathbf{k}), \tag{2.64}$$

$$= 1 - \frac{k_B T}{4\pi V \langle P^2\rangle_{NVE,E}}, \tag{2.65}$$

$$c_p = \frac{1}{N}\left(\frac{\partial H}{\partial T}\right)_p = \frac{V}{n k_B T^2} \lim_{\mathbf{k}\to 0} \langle \widehat{q}(\mathbf{k})\widehat{q}(-\mathbf{k})\rangle_{NVE} = \lim_{\mathbf{k}\to 0} c_p(\mathbf{k}), \tag{2.66}$$

$$= \frac{1}{N k_B T^2} \left\langle \Delta H^2 \right\rangle_{NPT}, \tag{2.67}$$

$$\kappa_T = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_T = \frac{V}{n^2 k_B T} \lim_{\mathbf{k}\to 0} \langle \widehat{n}(\mathbf{k})\widehat{n}(-\mathbf{k})\rangle_{NVE} = \lim_{\mathbf{k}\to 0} \kappa_T(\mathbf{k}), \tag{2.68}$$

$$= \frac{1}{V k_B T} \left\langle \Delta V^2 \right\rangle_{NPT} \tag{2.69}$$

where ρ, q and, n are the charge, heat and number densities, respectively, P is the macroscopic dipole, N is the number of particles, and, "NVE, **E**" and "NVE, **D**" indicate the microcanonical ensembles where the internal electric field and the applied field are fixed, respectively. The results of the simulations are displayed in Fig. 2.10, showing satisfactory agreement between Bayesian long-wavelength extrapolations, Eqs. (2.64) (2.66) (2.68), and standard fluctuation methods, Eqs. (2.65) (2.67) (2.69) [14, 74, 75].

Responses to homogeneous perturbations are not compatible with PBC but they can be estimated as the long wavelength limits of wave-vector dependent correlation functions. As proved in Fig. 2.10, the Bayesian regression protocol allows to evaluate the $\mathbf{k} \to 0$ limit and its error bar efficiently.

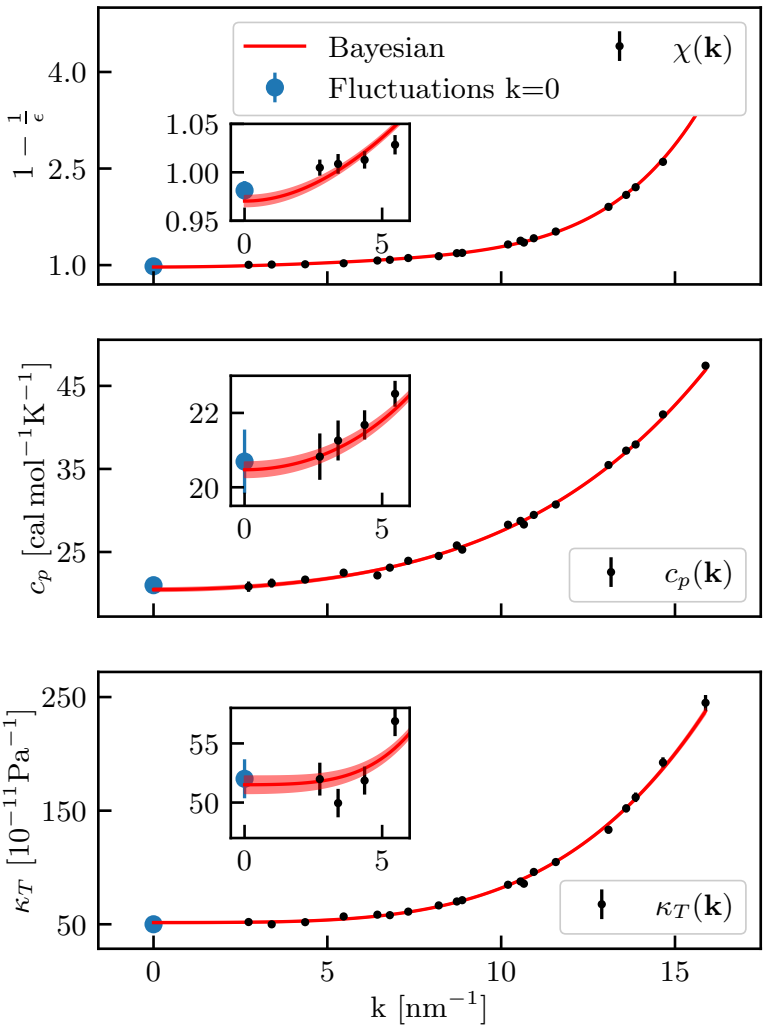

Figure 2.10: The black dots are estimates of isobaric specific heat, c_p, isothermal compressibility, κ_T, and dielectric constant, ϵ, computed through Eqs. (2.64) (2.66) (2.68), of liquid `TIP4P/2005` water at 350 K and 0 bar. The solid red line and the shaded red area are the Bayesian regression analyses and their statistical errors, fitted on Eqs. (2.64) (2.66) (2.68). The blue dots are the results of the fluctuation formulas, Eqs. (2.65) (2.67) (2.69).

3

SEEBECK COEFFICIENT OF CLASSICAL FLUIDS

In this chapter, I address thermoelectric effects in classical fluids. The thermopolarization coefficient of liquid water is studied in the framework of static linear response theory and computed via a tailored Bayesian linear regression scheme. In order to estimate the Seebeck coefficient in ionic conductors, we develop a Bayesian protocol based on the statistical properties of the thermoelectric Green Kubo estimator. Finally, we extend the Bayesian strategy proposing a comprehensive statistical model able to address the entire Onsager matrix.

3.1 THERMOPOLARIZATION EFFECT IN INSULATING POLAR FLUIDS

3.1.1 *Thermopolarization effect from static response theory*

Thermopolarization is of great theoretical and experimental interest [40–42, 75–85]. The response in polarization to a thermal field can be studied via EMD or NEMD simulations. Applying a thermal gradient in a simulation cell is a delicate procedure. In most calculations, a sink and source heat region, thermostatted at two different temperatures, are identified at the edges and in the middle of the simulation box, respectively [40–42, 76–82]. The temperature difference between the center and the edges must be small so that second and higher order responses are negligible, but large enough to observe a non zero signal. For these reasons, estimating the statistical uncertainty in a NEMD run is challenging. A static approach, based on linear response theory, for the computation of the thermopolarization coefficient is possible [75]. The Seebeck effect describes the electric field induced by a temperature gradient when no electric current flows [1, 21, 86, 87]:

$$\begin{aligned} \boldsymbol{J} &= \sigma \boldsymbol{E} - K_{12}/T \nabla T \\ \boldsymbol{J}_q &= K_{12} e \boldsymbol{E} - L_{qq} \nabla T, \end{aligned} \tag{3.1}$$

$$S \equiv \left. \frac{E}{\nabla T} \right|_{\boldsymbol{J}=0}. \tag{3.2}$$

where E and ∇T are any Cartesian components of the electric field and temperature gradient, $\boldsymbol{J}$ is the electric current, $\boldsymbol{J}_q$ is the heat current, e is the electron charge and $\{\sigma; K_{12}; L_{qq}\}$ are the Onsager coefficients as defined in Ref. [86]. The Seebeck coefficient is the ratio of one off-diagonal and one diagonal Onsager coefficient:

$$S = \frac{K_{12}}{\sigma T}. \tag{3.3}$$

In polar insulating fluids, an even in solid insulators, the ratio is an indeterminate form $\frac{0}{0}$. In these systems one cannot compute the Seebeck coefficient from the Onsager matrix, thus a new method is needed. An homogeneous temperature gradient is incompatible with PBC enforced in EMD simulations. One has to consider only periodic thermal perturbations commensurate with PBC. The equivalent mechanical perturbation that mimics a temperature inhomogeneity must induce the same unbalance in the heat density, q, defined in Eq. (1.62), [12]

$$\widehat{\mathcal{H}}' = -\frac{V}{T}\widehat{q}(\mathbf{k})T'(-\mathbf{k}), \tag{3.4}$$

where $T'(\mathbf{r}) = T(\mathbf{r}) - T$ is the deviation from the average temperature. Thanks to Gauss law, the Fourier transform of the polarization is $\boldsymbol{P}(\mathbf{k}) = i\mathbf{k}\rho(\mathbf{k})/k^2$. Analogously, the Fourier transform of the temperature gradient reads: $i\mathbf{k}T(\mathbf{k})$. In absence of external charges, the total field coincides with the polarization one. Therefore, for the Seebeck coefficient, one has:

$$S = \lim_{\mathbf{k}\to 0} S(\mathbf{k}), \text{ where} \tag{3.5}$$

$$S(\mathbf{k}) = -\frac{4\pi}{k^2}\chi_{q\rho}(\mathbf{k}), \tag{3.6}$$

$\chi_{q\rho}$ being the charge-temperature susceptibility. Thanks to static linear response theory, the charge-temperature susceptibility is the equal-time correlation function between heat and charge density:

$$\begin{aligned}\chi_{\rho q}(\mathbf{k}) &\equiv \frac{\partial\rho(\mathbf{k})}{\partial T(\mathbf{k})} \\ &= \frac{V}{k_B T^2}\left\langle \widehat{q}(\mathbf{k})\widehat{\rho}(-\mathbf{k})\right\rangle_0 .\end{aligned} \tag{3.7}$$

The heat density depends on the microscopic decomposition of energy into local contributions. The energy density can be redefined by adding the divergence of a bounded vector, $\boldsymbol{p}$: $e(\mathbf{r}) \to e(\mathbf{r}) + \nabla\cdot\boldsymbol{p}(\mathbf{r})$ [19, 44, 88]. Under the assumption that $\boldsymbol{p}$ is the gradient of a scalar field, the $\mathbf{k}\to 0$ limit of Eq. (3.5) is unaffected by this gauge freedom [75].

3.1.2 *Bayesian regression analysis of the thermopolarization coefficient of liquid water*

The static linear response approach is tested by computing the thermopolarization coefficient of liquid water at 400 K and 0 bar via Eq. (3.5) (3.6) and comparing the results to state-of-the-art NEMD calculations [76]. I simulate liquid water in a 0.5 ns-long EMD trajectory with 512 `SPC/E` water molecules [75, 89]. The `SPC/E` model assumes fixed OH bond distance and HOH angle, and assigns fix-point charges to the atoms. The equations of motions are integrated via the velocity-Verlet algorithm with a time-step of 0.25 fs, as implemented in `LAMMPS`[50].

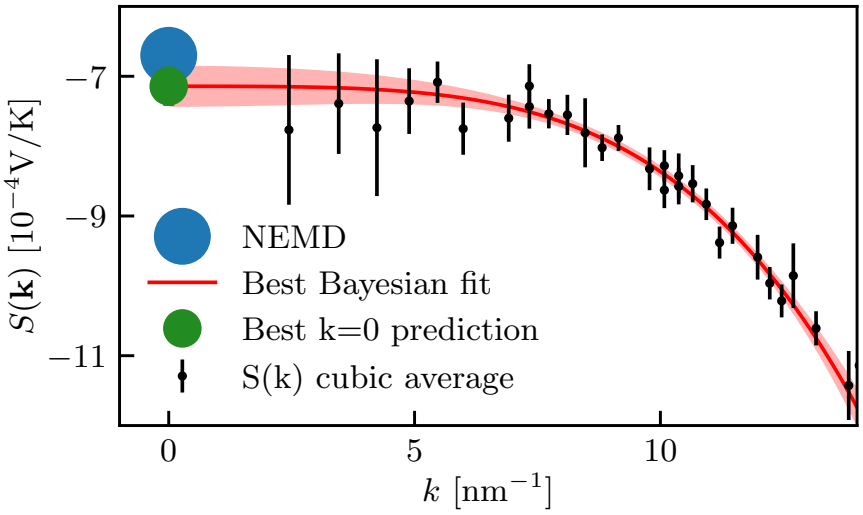

Figure 3.1: Wave-vector dependence of the Seebeck coefficient, $S(\mathbf{k})$, Eq. (3.6), of SPC/E water at $T = 400$ K and 0 bar [75]. The data are averages over equivalent wave-vectors. The red line is our Bayesian polynomial fit and the shaded area indicates the predicted uncertainty. At $\mathbf{k} = 0$ I report the extrapolation and the value obtained by Wirnsberger et al. [76] via NEMD. Figure reproduced from Ref. [75].

The samples of the wave-vector dependent Seebeck coefficient, Eq. (3.7), are averages computed over the trajectory and therefore are Gaussian variables whose variance decreases increasing the trajectory's length. In order to extrapolate the long-wavelength limit in Eq. (3.5), I employ the Bayesian extrapolation method detailed in Sec. 2.4. I approximate Eq. (3.6) with a low order polynomial, defined in Eq. (2.52),

$$S(\mathbf{k}) \approx w_0 + w_1 k^2 + w_2 k^4 \cdots + w_P k^{2P}.$$

The procedures for prior optimization and model selection are described in Sec. 2.4. The outcomes of Eq. (3.6) and the Bayesian linear regression results are displayed in Fig. 3.1. The long-wavelength limit, Eq. (3.5), is in good agreement with previous state-of-the-art NEMD calculations [76]. In Fig. 3.2, I display the optimal model selection procedure by showing the probability of the number of parameters after optimizing the prior distribution, discussed in Eq. (2.60) [61, 75]. Fig. 3.2 shows that the model selection protocol is crucial to prevent over-fitting. Considering a dataset containing 29 elements, the Bayesian analysis selects the model with 3 parameters.

The optimal number of parameters and the $\lim_{k\to 0} S(\mathbf{k})$ prediction are dependent on the maximum $\mathbf{k}$-vector included in the Bayesian regression. Under the assumption that the wave-vectors are sufficiently small in magnitude, the long-wavelength extrapolation and the optimal number of parameters should converge as one enlarges the dataset on which the regression is performed. Fig. 3.3 shows the convergence of the prediction and of the selected optimal

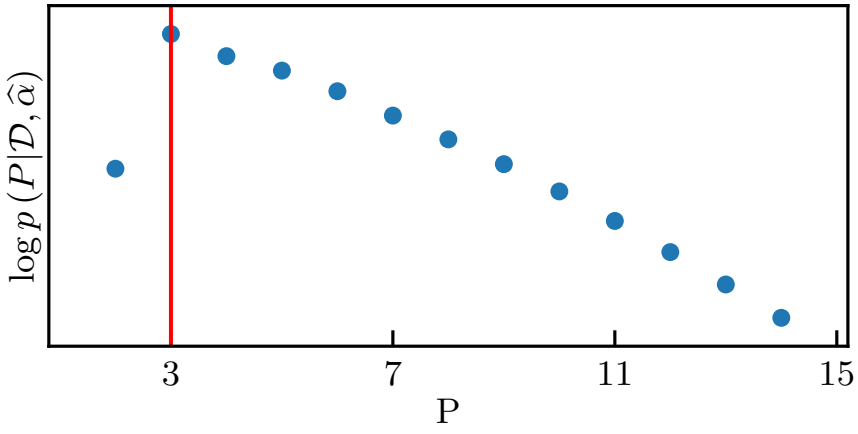

Figure 3.2: Logarithm of the probability distribution of the number of parameter upon optimization of the Gaussian prior hyperparameter α, defined in Eq. (2.60). The vertical red solid line indicates the maximum probability as a function of the number of parameters of the polynomial fit.

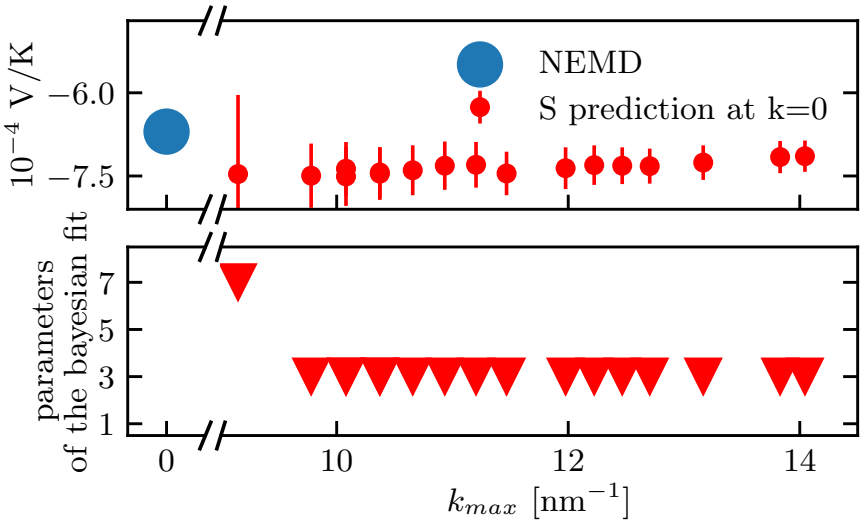

Figure 3.3: Upper panel: Bayesian linear regression prediction of the $\mathbf{k} = 0$ value of the Seebeck coefficient of SPC/E water at T = 400 K and 0 bar, as a function of the maximum magnitude of the $\mathbf{k}$-vectors, k_{max} [75]. The NEMD result from Ref. [76] is also reported. Lower panel: Optimal number of parameters of the Bayesian regression fit as a function of k_{max}. Figure reproduced from Ref. [75].

number of parameters, as a function of the $\mathbf{k}$-vector with maximum magnitude, k_{max}, included in the fit. The outcomes prove that our method is numerically stable.

The size effects of Eq. (3.6) are estimated computing the Seebeck coefficient from a 0.5 ns-long simulation with 1728 SPC/E water molecules at 400 K and 0 bar. Fig. 3.4 presents the outcomes of the investigation of larger box sizes,

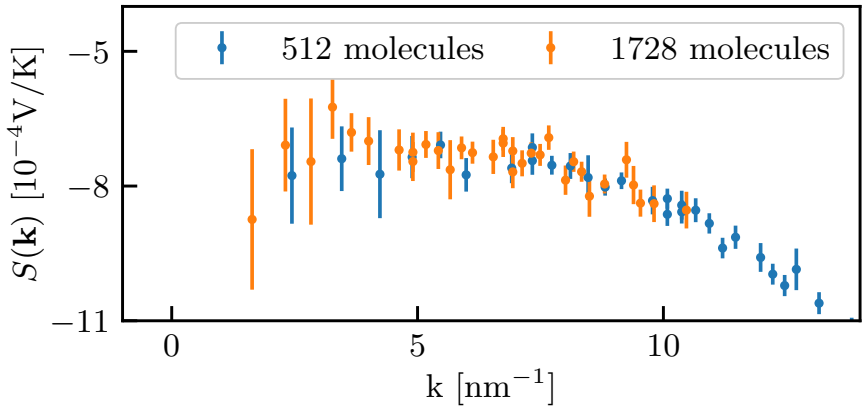

Figure 3.4: Wave-vector dependence of the Seebeck coefficient, $S(\mathbf{k})$, Eq. (3.6), of 512 and 1728 SPC/E water molecules at T = 400 K and 0 bar [75]. The data are averages over equivalent wave-vectors.

highlighting that the statistical uncertainty of Eq. (3.6) increases when probing wave-vectors of smaller magnitudes. Additionally, the analysis does not report significant size effects. Since numerical estimates of Eq. (3.6) at $\mathbf{k}$-vectors with small magnitudes have larger error bars, the brute force extrapolation of the long-wavelength limit —simulating large systems— would have been challenging. Bayesian regression allows to estimate the long-wavelength limit and its statistical uncertainty from small box-sizes and short EMD simulations, reducing the computational burden of the numerical method.

Since the Bayesian procedure is stable and accurate for the extrapolation of long-wavelength limits of periodic static response functions, I investigate the Seebeck coefficient of liquid SPC/E water in a wide temperature-pressure range and compare our results with previous NEMD calculations [75, 80]. The long-wavelength extrapolations of the Seebeck coefficient, reported in Fig. 3.5, are in good agreement with NEMD results [75, 80].

In this section I proved that the polarization induced by a temperature gradient is a quasi-equilibrium effect [75]. In a polar insulating fluid, after being perturbed by a temperature gradient, the dielectric degrees of freedom equilibrate in a microscopic time scale, compared to the one of hydrodynamic variables, thus allowing to apply static linear response theory to thermopolarization effects. I benchmarked the static approach against state-of-the-art NEMD simulations of water [76, 77], showing that the protocol is valid and predictive. Leveraging on Bayesian linear regression methods, I estimated the long-wavelength limit in Eq. (3.5), overcoming some of the numerical difficulties of large-size extrapolations [75].

In conclusion, on the applicative side, in the presence of large temperature gradients, water can display huge polarization, especially if the electric susceptibility of the fluid is enhanced by complex polar molecules. This conditions has been recently reported in mitochondria, the furnaces of the cell [90]. Indeed, a temperature drop of about 10 °C was allegedly observed across the

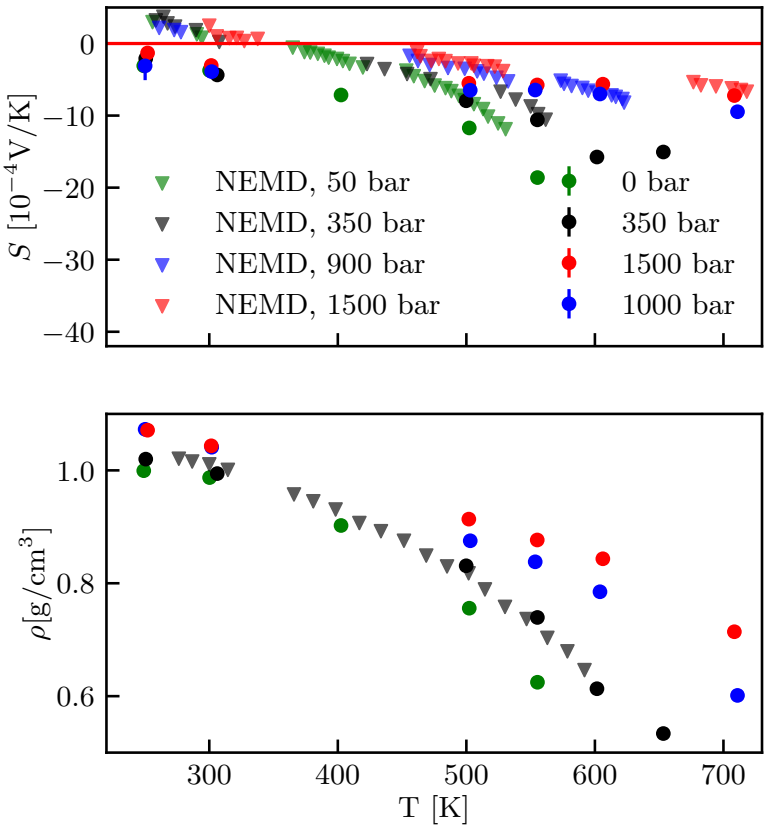

Figure 3.5: Upper panel: Seebeck coefficient, S, Eq. (3.5), estimated at various temperature-pressure conditions, via Bayesian linear regression [75]. The results are compared with NEMD simulations in Armstrong et al. [80]. Lower panel: Mass density of SPC/E water evaluated at different temperature-pressure conditions and compared with Armstrong et al. [80] at 350 bar [75]. Figure reproduced from Ref. [75].

intermembrane space whose extent is of the order of 10 nm [90]. Even if the claim has been mitigated by others [91, 92], cellular metabolism requires huge amounts of energy and may be the source of large thermal gradients [90–92]. The numerical simulations of SPC/E water suggest a electromotive force of 2–6 mV generated by a 4–12 °C temperature drop across a 12 nm slab. In order to draw definitive conclusions on the thermoelectric properties of mitochondria, too little is known about the temperature distribution in the cell and on the composition of the intermembrane fluid. Further investigations might shed some light on the thermoelectric interplay in the cell metabolism mechanism, which could open unexplored avenues for molecular simulations in life science [75].

3.2 THERMOELECTRIC EFFECTS IN IONIC CONDUCTORS

3.2.1 *Seebeck effect and thermoelectric transport in ionic conductors*

In Sec. 3.1, I addressed the polarization induced by a temperature gradient in polar fluids. In such systems, charge transport is forbidden. Conversely, in ionic conductors charged ions diffuse. In insulators, the Seebeck effect is a quasi-equilibrium phenomenon, whereas in conducting systems, it is related to charge transport. In ionic conductors, when a temperature gradient, or an electric field, is applied, charge transport occurs. The Onsager equations in Eq. (3.1) set a linear relation between the external thermodynamic forces and the induced currents:

$$\begin{aligned}\boldsymbol{J} &= \sigma\boldsymbol{E} - \mathsf{K}_{12}/\mathsf{T}\nabla\mathsf{T}\\ \boldsymbol{J}_q &= \mathsf{K}_{12}e\boldsymbol{E} - \mathsf{L}_{qq}\nabla\mathsf{T},\end{aligned}$$

In Eq. (1.60), the heat current, $\boldsymbol{J}_q$, depends on the partial enthalpies, h_i: $\boldsymbol{J}_q = \boldsymbol{J}_e - \sum_i h_i \boldsymbol{J}_i$. The method used for the calculation of the partial enthalpies is based on the discussion presented in Refs. [22–24] and is reviewed in Appendix C [21]. I am interested in thermoelectric phenomena, such as the application of an electric field and the corresponding rise of the heat current, or viceversa, of a temperature gradient and the resulting diffusion of ions. The Seebeck effect describes the electric field generated by a temperature gradient in absence of an electric current. Depending on open or closed circuit conditions, the Seebeck effect refers to completely different processes [93, 94]. After a temperature gradient is applied, a conducting closed system (open circuit) initially transports charge. In the long-time limit the system polarizes, and eventually, no electric current flows. In a open system (closed circuit), a steady electric current rises. To achieve zero net current, I must impose an electric field to counterbalance the electric flow induced by the temperature gradient. Here, I study closed circuit conditions: the Seebeck effect refers to the electric field that must be applied to an open system, perturbed by a temperature gradient, in order to obtain zero net electric currents [93]. The Seebeck coefficient, Eqs. (3.2) (3.3), is defined as

$$\begin{aligned}\mathsf{S} &\equiv \left.\frac{\mathsf{E}}{\nabla\mathsf{T}}\right|_{\boldsymbol{J}=0}\\ &= \frac{\mathsf{K}_{12}}{\sigma\mathsf{T}}.\end{aligned}$$

Although NEMD simulations directly implement thermodynamic forces, they require careful checks to avoid size effects and, second and higher order responses [43, 95–97]. In this section I analyze the statistical and numerical challenges of the calculation of the Seebeck coefficient through EMD simula-

tions. The GK theory provides a solid theoretical background for computing Onsager coefficients, Eqs. (1.74) (1.77) (2.8),

$$\begin{aligned} \sigma &= \frac{V}{3k_B T}\int_0^\infty \left\langle \widehat{\boldsymbol{J}}(t)\cdot\widehat{\boldsymbol{J}}(0)\right\rangle_0 dt\,, \\ K_{12} &= \frac{V}{3k_B T^2}\int_0^\infty \left\langle \widehat{\boldsymbol{J}}(t)\cdot\widehat{\boldsymbol{J}}_q(0)\right\rangle_0 dt\,, \\ L_{qq} &= \frac{V}{3k_B T}\int_0^\infty \left\langle \widehat{\boldsymbol{J}}_q(t)\cdot\widehat{\boldsymbol{J}}_q(0)\right\rangle_0 dt\,. \end{aligned} \tag{3.8}$$

As discussed in Sec. 2.1, the numerical evaluation of Onsager coefficients via the GK integrals is far from trivial. The block analysis of HE and GK integrals requires long EMD trajectories to reach convergence and reasonable error bars. Thanks to the Wiener-Khintchine theorem [51, 52], the GK integrals can be rewritten as the 0-frequency limit of the PSD, Eq. (2.30),

$$S^{ij}(\omega) = \int_0^\infty \left\langle \widehat{\boldsymbol{J}}^i(t)\cdot\widehat{\boldsymbol{J}}^j(0)\right\rangle e^{i\omega t}\,dt\,,$$

where $\boldsymbol{J}^i$ can be any current entering the Onsager equations, in the present case $\boldsymbol{J}^i = \{\boldsymbol{J}, \boldsymbol{J}_q\}$.

3.2.2 *Bayesian regression analysis of the Seebeck coefficient*

We apply the tools developed in Sec. 2.3 to compute the Seebeck coefficient through GK theory [46]. The off-diagonal elements of the periodogram matrix, Eq. (2.33), are distributed according to the Gamma-variance distribution, whose probability density, defined in Eq.(2.34), is

$$\begin{aligned} p\left(S_k^{cq}\right) = &\frac{\left|S_k^{cq}\right|^{\frac{\nu-1}{2}}}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{2^{\nu-1}\pi(1-\rho(\omega_k)^2)}[\sigma(\omega_k)L_{qq}(\omega_k)]^{\frac{\nu+1}{4}}} \\ &\times \mathcal{K}_{\frac{\nu-1}{2}}\left(\frac{\left|S_k^{cq}\right|}{[\sigma(\omega_k)L_{qq}(\omega_k)]^{1/2}(1-\rho(\omega_k)^2)}\right) \\ &\times \exp\left\{\frac{\rho(\omega_k)S_k^{cq}}{[\sigma(\omega_k)L_{qq}(\omega_k)]^{1/2}(1-\rho(\omega_k)^2)}\right\}, \end{aligned} \tag{3.9}$$

where S^{cq} is the charge-heat component of the periodogram. $\sigma(\omega)$ and $L_{qq}(\omega)$ are the PSD of the charge and heat fluxes, reported in Eq. (3.8). The correlation coefficient estimator is

$$\widehat{\rho}_k = \widehat{S}_k^{cq}(\sigma(\omega_k)L_{qq}(\omega_k))^{-1/2}. \tag{3.10}$$

$\widehat{\rho}_k$ is independent across different frequencies and its probability density is reported in Eq. (2.36).

We define a Bayesian regression analysis to filter $\widehat{\rho}_k$. The model of $\rho(\omega)$, defined in Eq. (2.36), is a cubic spline, Θ_{cs} in Eq. (2.49), $\rho(\omega) \approx \Theta_{cs}(\omega|\boldsymbol{\theta})$, [46, 63]. The parameters of the model, $\boldsymbol{\theta} = \{\theta_0, \ldots, \theta_{P-1}\}$, are optimized via Bayesian

regression as explained in Sec. 2.3. We identify Eq. (2.36) with the Likelihood distribution of the Bayesian regression procedure [46]. Since correlation coefficients are independent across different frequencies, the Likelihood distribution reads $p(\{\rho_k\} \mid \boldsymbol{\theta}) = \prod_k p(\rho_k \mid \boldsymbol{\theta})$. The expectation value of the parameters over the posterior distribution, defined in Eq. (2.41), is estimated as a MCMC average implemented in the `emcee` code [64], assuming a uniform prior. The convergence of the MCMC is carefully checked analyzing the autocorrelation time of the Monte Carlo realizations, τ_{MCMC}, making sure that the length of the chain, $\mathcal{N}$, is $\mathcal{N} > 100\tau_{\text{MCMC}}$ [46]. The Seebeck coefficient reads

$$S = \frac{1}{T} \left\langle \widehat{\theta}_0 \right\rangle_{p(\boldsymbol{\theta}|\{\rho_k\})} \sqrt{\frac{L_{qq}(0)}{\sigma(0)}}. \tag{3.11}$$

The optimal number of parameters of the spline model, Θ_{cs} in Eq. (2.49), is selected via AIC, Eq. (2.26),

$$\text{AIC}(P) = -2 \max_{\boldsymbol{\theta}} \log\left[p(\{\rho_k\} \mid \boldsymbol{\theta}, P)\right] + 2P. \tag{3.12}$$

As in Ref. [46], $L_{qq}(0)$ and $\sigma(0)$ are computed via cepstral analysis, described in Sec. 2.2. Since the correlation coefficient is a even function of the frequency, a model with P parameters corresponds to a spline with $2P - 1$ knots.

We compute the Seebeck coefficient of molten CsF, NaCl, KCl, LiCl at different pressure-temperature conditions from EMD simulations performed using `LAMMPS` [46, 50]. The CsF system is modeled with a sample of 512 atoms interacting through the Born-Mayer-Huggins-Tosi-Fumi force field [47–49]. We equilibrate the systems in the NpT ensemble for 200 ps and in the NVT ensemble for 200 ps [46]. The fluxes are harvested from NVE trajectories whose length ranges from 40 to 400 ns. Energy and charge fluxes are sampled every fs.

As a benchmark, we estimate the Seebeck coefficient of molten CsF at 1400 K and 0 bar via the GK formula from a 40 ns long trajectory [46]. The reference GK result is estimated via block analysis over 1 ns-long segments. In Fig. 3.6, we compare, as a function of the length of the trajectory, τ, the Seebeck coefficient computed via Bayesian regression and block analysis [46]. The Bayesian estimate converges rapidly to the reference one. The variance of both methods is displayed in Fig. 3.7, which proves that Bayesian regression analysis yields an efficient estimator of the Seebeck coefficient, whose variance decreases increasing τ. Compared to block analysis, Bayesian regression is almost one order of magnitude more efficient, allowing trajectory lengths to be limited to a few nanoseconds.

To further investigate the statistical properties of the Bayesian regression estimator, we analyze the distribution of the Bayesian regression results obtained from 200 2 ns-long EMD simulations of molten CsF at 1200 K and 0 bar [46]. In Fig. 3.8 we present the distribution of the outcomes: the solid red line is a Gaussian curve fitted to the raw data while the solid green line is a Gaussian centered on the average of the predictions with its width estimated from the average of the predicted statistical uncertainties [46]. The two curves are

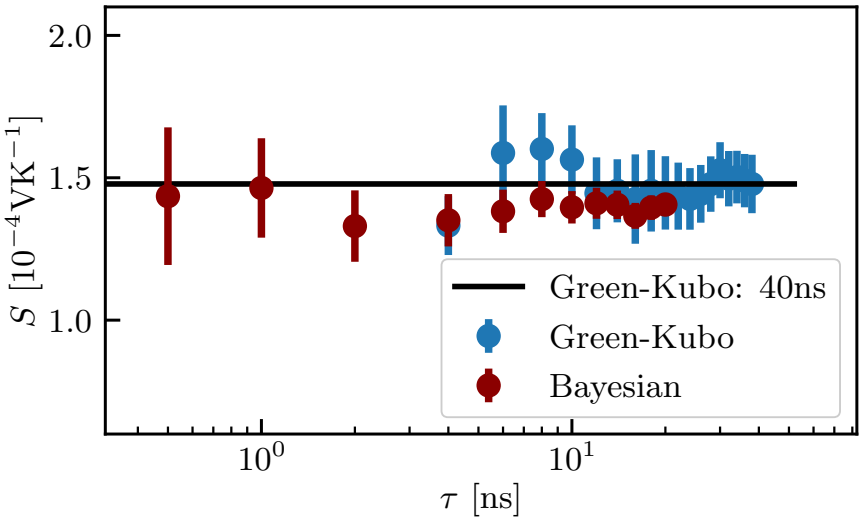

Figure 3.6: Seebeck coefficient of CsF at 1400 K and 0 bar computed from GK and the Bayesian regression as a function of the length of the equilibrium MD trajectory.

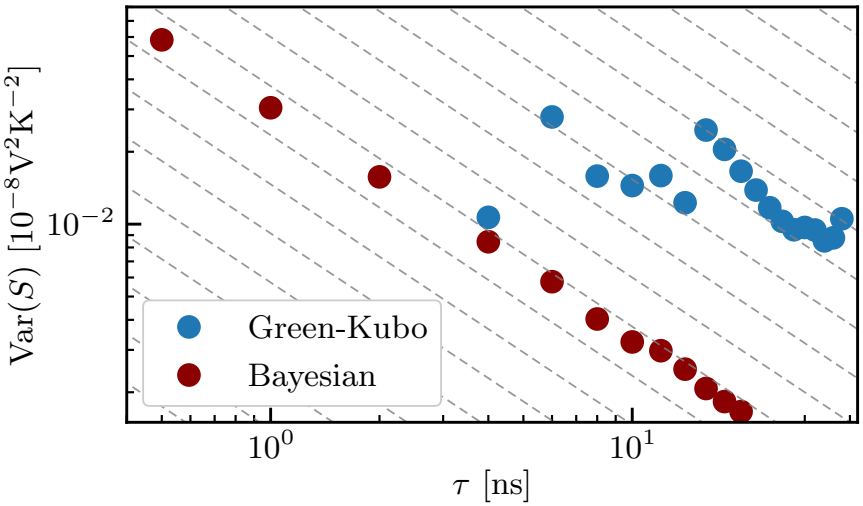

Figure 3.7: Variance of the Seebeck coefficient of CsF at 1400 K and 0 bar computed from GK and the Bayesian regression as a function of the length of the equilibrium MD trajectory. The dashed line indicates the τ^{-1} scaling.

in good agreement, showing that the Bayesian standard deviation is compatible with the true one. The Bayesian regression analysis outcomes pass the Shapiro-Wilk normality test with a level of significance of 0.05 [58].

By properly modifying the Likelihood function, Bayesian regression can address diagonal elements of the periodogram matrix, defined in Eq. (2.14) [46]. In Fig. 3.9, we display the periodogram of the electric current, defined in Eqs. (2.9) (2.14), of molten CsF at 1400 K and 0 bar, obtained from a 5 ns EMD simulation [46]. The Bayesian regression analysis is in good agreement with

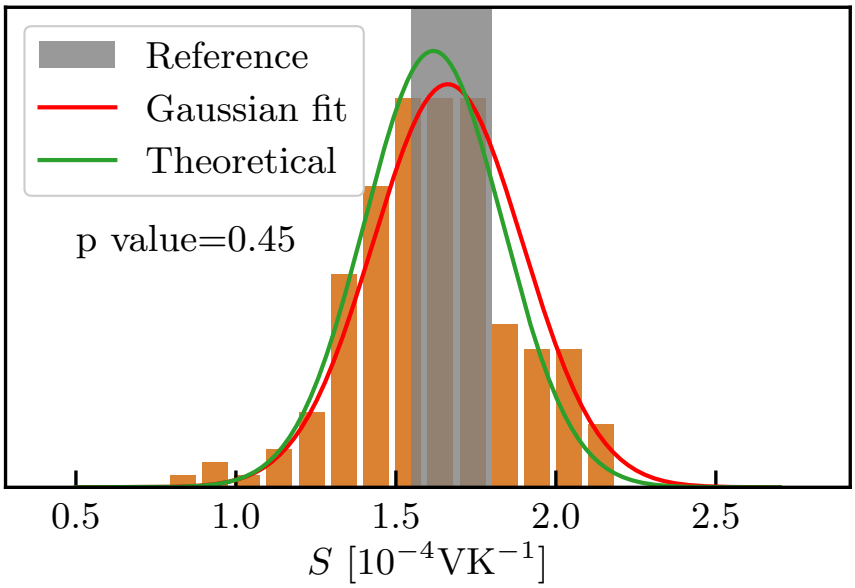

Figure 3.8: Distribution of the Bayesian estimates of the Seebeck coefficient extracted from 200 2 ns-long trajectories of CsF at 1200 K and 0 bar, and the p-value of the Shapiro-Wilk normality test. The solid red line is a Gaussian curve fitted on the distribution of the Bayesian predictions while the solid green line is a Gaussian curve centered on the average of the predictions with, as standard deviation, the average of the uncertainty estimated from the Bayesian analysis. The vertical gray band represent the reference GK estimate of the Seebeck coefficient estimated via block analysis from a 40 ns-long simulation [46]. Figure reproduced from Ref. [46].

the cepstral filter on the same data. Compared to the virtually negligible cost of cepstral analysis, implementing MCMC significantly impacts the computational efficiency of the Bayesian method.

After careful benchmarks on CsF, we extend the study to molten NaCl, KCl, and LiCl under several pressure-temperature conditions [46]. NaCl, KCl, and LiCl are simulated with a sample of 1000 atoms modeled with the Deep Potential Molecular Dynamics (DeePMD) neural network interatomic potential [99–101]. The Hamiltonian equations of motion are integrated using the Velocity-Verlet algorithm with a 1 fs time step, as implemented in `LAMMPS` [50]. We equilibrate the systems in the NpT ensemble for 500 ps and in the NVT ensemble for 200 ps. After equilibration, energy [28] and electric currents are harvested from a 2 ns NVT [102] trajectory sampled every 2.5 fs [46]. Machine learning potentials do not feature explicit atomic charges. Thanks to the gauge invariance principle of transport coefficients, integer oxidation charges can be assigned to the atomic species to estimate Onsager coefficients [19, 103, 104]. Fig. 3.10 displays the Seebeck coefficient of molten CsF, NaCl, KCl, LiCl at several temperature-pressure conditions. The Bayesian regression analysis on NaCl and KCl shows order of magnitude-agreement with experimental measures of thermopower reported in Ref. [105]. It is interesting to notice that the

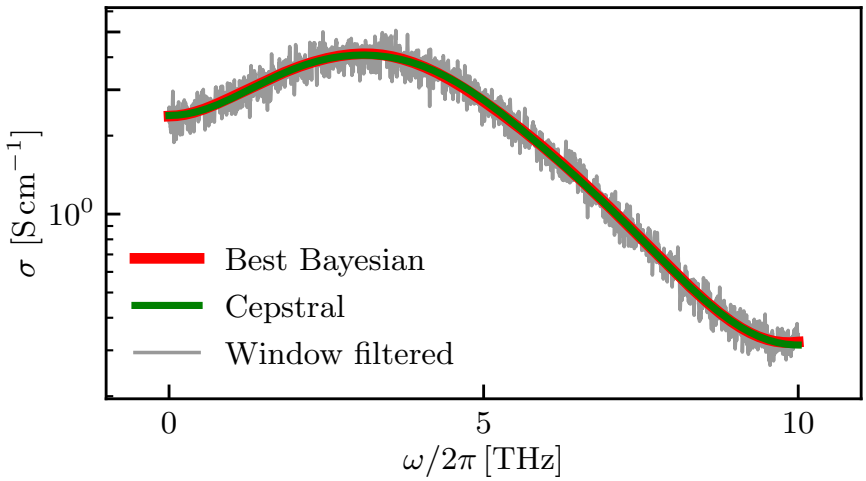

Figure 3.9: Best Bayesian regression analysis of the electric conductivity of molten CsF at 1400 K and 0 bar as a function of frequency, obtained from a 5 ns equilibrium MD simulation. The best model has been selected via the AIC. The green line is the cepstral-filtered spectrum. The window-filtered line refers to a moving average [98] performed with a window width of 0.01 THz.

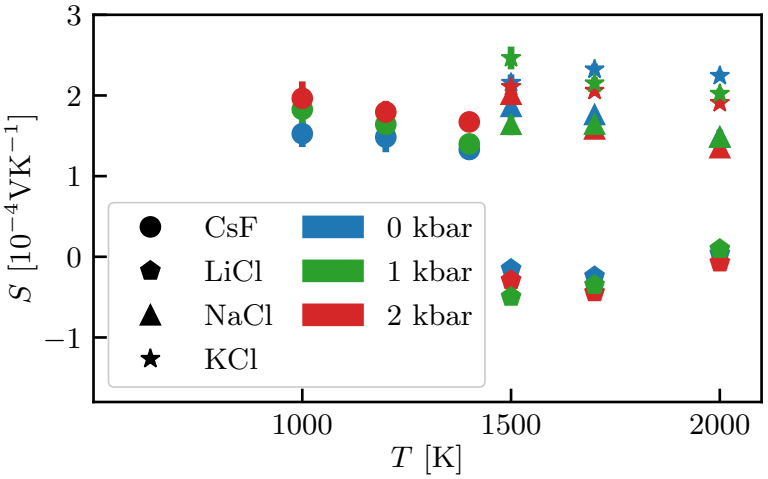

Figure 3.10: Seebeck coefficient of molten LiCl, CsF, NaCl and KCl as functions of pressure and temperature [46]. The Seebeck coefficient has been estimated via the Bayesian regression analysis as detailed in the text. Figure reproduced from Ref. [46].

Seebeck coefficient change sign between LiCl and the other molten salts. Indeed, an isotopic effect has been reported to affect the Seebeck coefficient [106]. In the case of LiCl and CsF, the change in sign is possibly due to the considerable inverted anion-cation mass ratio.

Via extensive MCMC sampling, the Bayesian regression protocol yields a consistent and accurate estimator of the Seebeck coefficient [46]. The procedure leverages Bayesian regression to tackle off-diagonal —as well as diagonal—

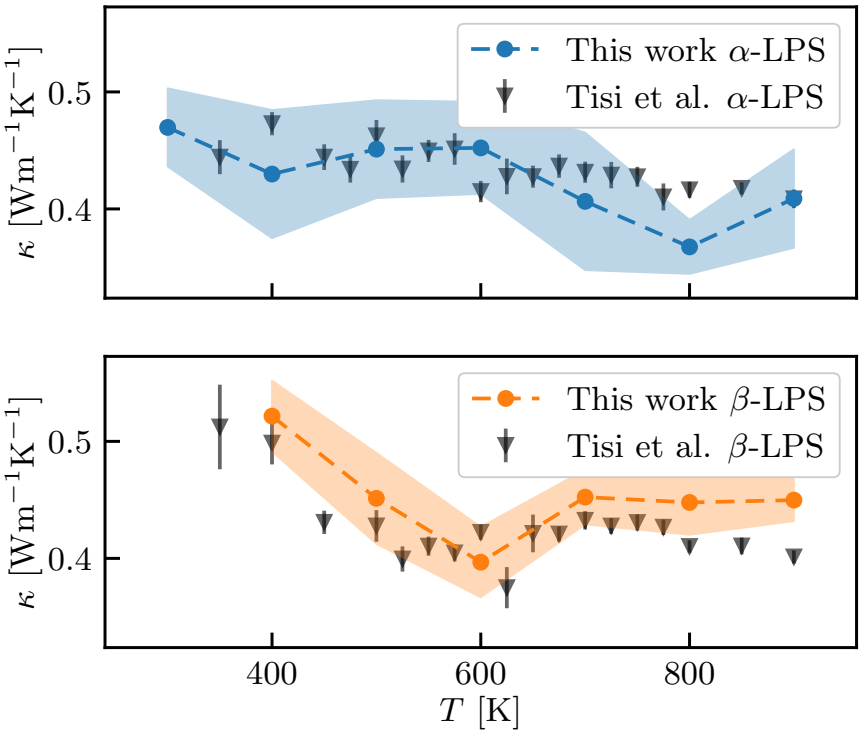

Figure 3.11: Heat conductivity of α (upper panel) and β (lower panel) LPS at different temperatures computed by NLL Bayesian regression analysis and compared to the results of Ref. [55].

transport coefficients from moderately-long EMD simulations. Being able to address the Seebeck effect, the Bayesian approach inaugurates a novel strategy to assess thermoelectric efficiency of ionic materials from relatively short EMD simulations [45, 46].

3.2.3 *Bayesian regression analysis of transport coefficients*

Spectral methods target the statistical properties of the periodogram matrix, defined in Eq. (2.33), enabling the efficient computation of transport coefficients from relatively short EMD simulations [44, 75]. A single statistical model can address the entire Onsager matrix in a all-encompassing Bayesian regression protocol [45]. Some of the numerical drawbacks of EMD methods described in Sec. 2.1, Sec. 2.2 and, Sec. 3.2.1 can be partially lifted by a comprehensive Bayesian regression analysis [45].

The probability density of a Wishart-distributed stochastic process, e. g. $\widehat{\mathbf{S}}_k$ defined in Eq. (2.33), is:

$$p(\mathbf{S}_k) = \frac{(\det\{\mathbf{S}_k\})^{(\ell-M-1)/2} e^{-\mathrm{Tr}\left(\mathbf{S}(\omega_k)^{-1}\mathbf{S}_k\right)/2}}{2^{\frac{\ell M}{2}} (\det \mathbf{S}(\omega_k))^{\ell/2} \Gamma_M(\frac{\ell}{2})}, \tag{3.13}$$

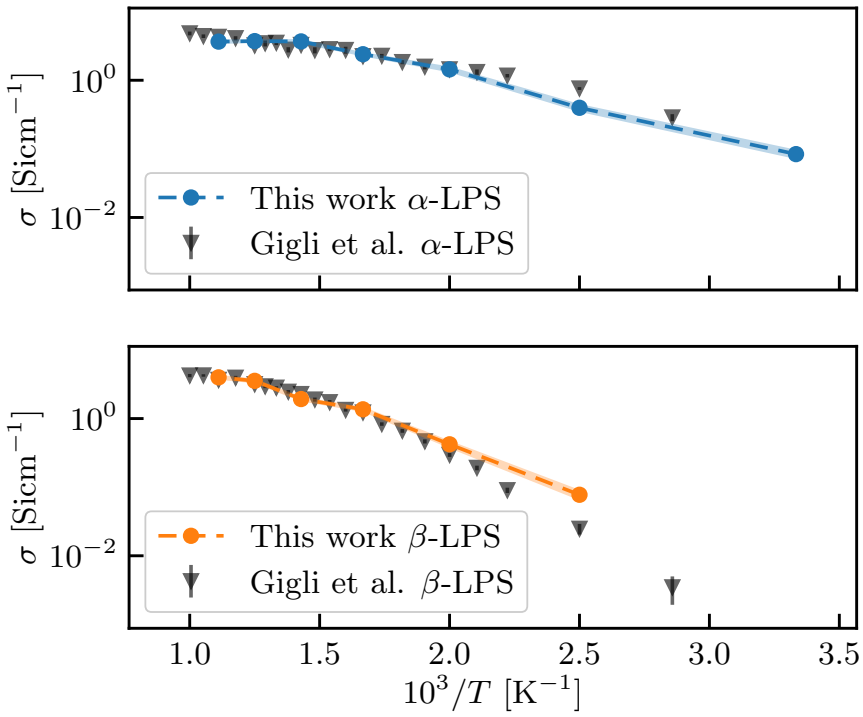

Figure 3.12: Electric conductivity of α (upper panel) and β (lower panel) LPS at different temperatures computed by NLL Bayesian regression analysis and compared to the results of Ref. [56].

where M is the number of thermodynamic fluxes, ℓ is the number of equivalent realizations of the currents, in a isotropic system ℓ equals 3, and $\mathbf{S}$ is a $M \times M$ positive-definite matrix.

One can design a single Bayesian model to infer the entire periodogram matrix [45]. For the Bayesian regression of the periodogram matrix, Eq. (3.13) is chosen as the Likelihood distribution. In order to impose positive definiteness, $\mathbf{S}(\omega)$ is rewritten as the product of a triangular matrix, $\mathbf{C}(\omega)$: $\mathbf{S}(\omega) = \mathbf{C}(\omega)\mathbf{C}(\omega)^{\mathsf{T}}$. We model each entry of $\mathbf{C}(\omega)$ using a spline with P knots, as defined in Eq. (2.49) [45]. Building on the discussion detailed in Sec. 2.3 and Sec. 3.2, we estimate the optimal value of the parameters from their posterior distribution, Eq. (2.41), assuming a uniform prior [45]. Because of the computational cost of the MCMC [46], here we evaluate the optimal value and the variance of the parameters minimizing the Negative Log-Likelihood (NLL) in the Laplace approximation [45, 107]. The minimization usually requires far fewer steps than the MCMC to achieve convergence. In the Laplace approximation, the covariance matrix of the parameters is derived expanding the NLL to second order [107]. The optimal number of knots of the spline is selected via AIC, as discussed in Eq. 3.12 and Sec. 2.2.

The NLL method is tested computing the transport coefficients of Li_3PS_4 (LPS). LPS is recognized as one of most promising compounds in the family of sulfide Solid State Electrolyte (SSE) [55, 56], in which the Li ions diffuse

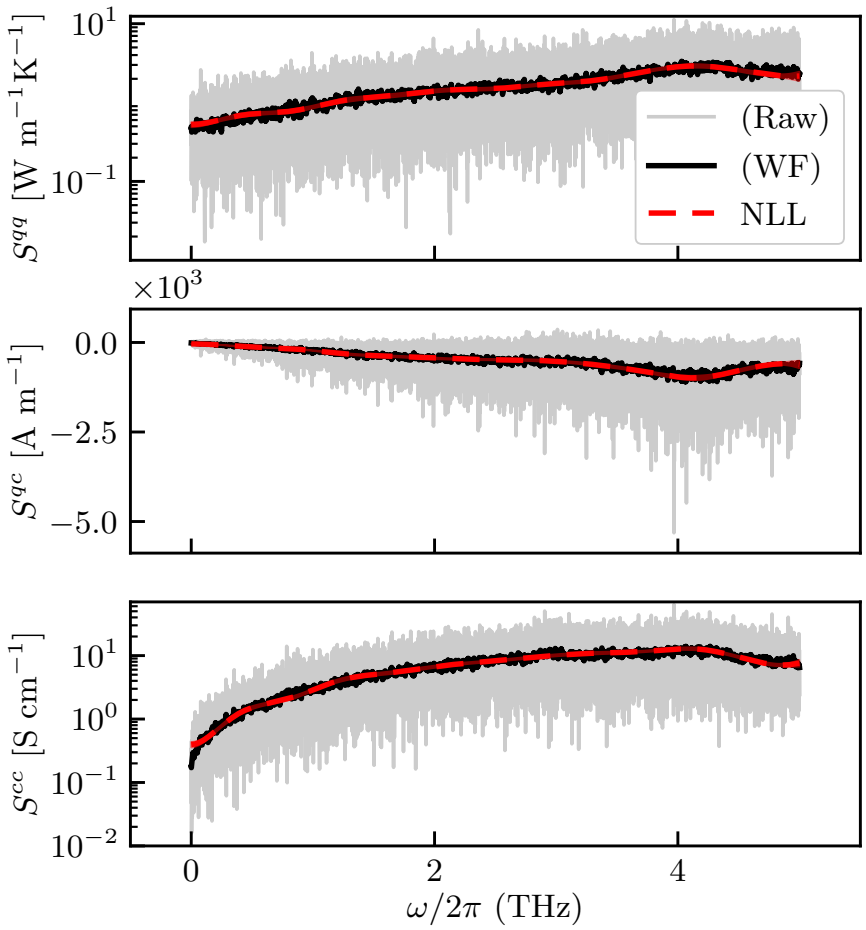

Figure 3.13: Heat (upper panel), heat-charge (middle panel) and charge (lower panel) periodogram computed from a 5 ns-long trajectory of α-LPS at 400 K. The black solid line indicates a WF of width 0.01 THz and the red dotted line is the NLL Bayesian regression analysis.

while the other atoms maintain a well-defined solid structure. LPS is present in three different phases: α, β and γ [108–113]. The inter-atomic potential of LPS is fitted via the machine learning neural network Neuroevolution Potential (NEP), as discussed in Appendix D. Even if the γ phase is the most stable at room temperature, it is the least conducting [56]. For this reason, we disregard the analysis of the γ phase. α and β-LPS are simulated in a wide range of temperatures [45, 56]. The trajectories are integrated via the velocity-Verlet algorithm as implemented in the `GPUMD` code with a 0.5 fs time step [114–116]. LPS is modeled with a sample of 6144 atoms and equilibrated in the NVT ensemble for 10 ns [102]. After equilibration, the production data are harvested every time-step from a 5 ns-long NVT trajectory [102]. The NEP force field is not able to predict atomic charges. Thanks to the gauge invariance of transport coefficients, we can assign oxidation charges to the atoms without affecting the Onsager matrix [19, 46, 104].

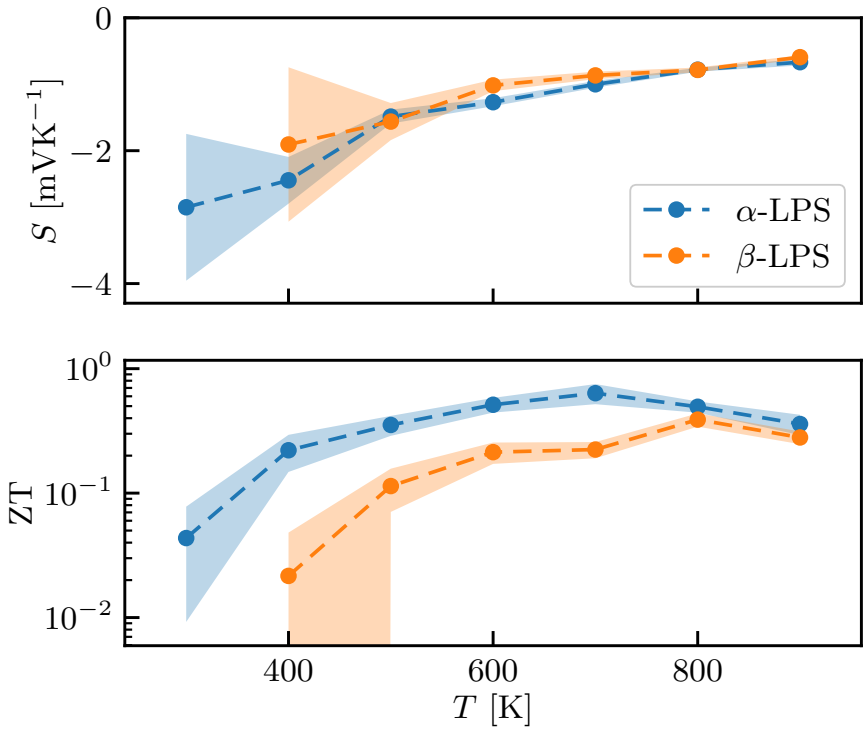

Figure 3.14: Seebeck coefficient (upper panel), figure of merit (lower panel) of α and β-LPS, computed via NLL Bayesian regression analysis, as a function of temperature.

In Fig 3.11 and Fig. 3.12, we validate the NEP by computing, via the NLL Bayesian regression method, the electric and heat conductivity of α and β-LPS in a wide range of temperatures [45]. The NLL results are in good accordance with previous calculations [55, 56]. The slight deviation visible in Fig 3.11-3.12, can be due to the smaller box sizes considered in Ref. [55, 56]. The NLL approach yields an accurate estimator of diagonal elements of the periodogram matrix and is computationally more efficient than the MCMC method discussed in Ref. [46]. The Bayesian regression of the entire periodogram matrix is based on the statistical properties of a Wishart process. The NLL analysis of the periodogram of α-LPS at 400 K is reported in Fig. 3.13. The results are in good agreement with the Window Filter (WF) applied to the raw data, proving that the spline model is expressive enough to capture the features of the periodogram.

The Seebeck coefficient and heat conductivity of a multi component system, Eqs. (3.3) (1.76), are not Onsager coefficients, but can be derived easily from the periodogram [15, 45, 46]. In Fig. 3.14 we display the Seebeck coefficient and figure of merit of α and β-LPS in a wide range of temperatures [45]. The figure of merit, $ZT = \frac{S^2\sigma}{\kappa}T$, is a measure of thermoelectric efficiency [117, 118]. ZT depends dramatically on S and its typical values range between 10^{-3} —poor thermoelectric material— to 1 —excellent thermoelectric material—.

α and β-LPS demonstrate competitive thermoelectric performance [117, 118]. The best efficiency is achieved at relatively high temperatures, which in part overshadows the good thermoelectric properties of LPS. Bayesian analysis is able to access the information stored across all frequencies of the periodogram. The Laplace approximation of the NLL reduces the computational cost of the optimization procedure compared to MCMC [45, 107]. A single statistical model filters the Onsager matrix, accounting for the different statistical properties of the periodograms.

4

THERMAL TRANSPORT IN SOLID AND LIQUID INSULATORS

In quasi equilibrium conditions, the hydrodynamic equations govern the relaxation of the densities of conserved extensive variables. In this chapter, I study heat propagation in solids and liquids. In 1994, Palmer proposed to extrapolate thermal conductivity through energy density fluctuations. We review this approach and overcome some of its numerical difficulties thanks to cepstral analysis. The energy density fluctuations method is compared to a recently proposed technique, which instead addresses mass density fluctuations. Next, I briefly overview the Quasi-Harmonic-Green-Kubo theory of heat transport in solids. Careful numerical simulations on large harmonic models of glasses reveal disorder-induced Rayleigh scattering at low frequencies. When considering only Rayleigh scattering in harmonic disordered systems, the predicted thermal conductivity diverges at all temperatures, irrespective of configuration disorder.

4.1 EXTRAPOLATION OF THERMAL CONDUCTIVITY FROM ENERGY DENSITY FLUCTUATIONS

In this section, I describe how to estimate transport coefficients employing density rather than current fluctuations. In 1994, Palmer proved that the regression of density fluctuations towards equilibrium is related to transport coefficients, which are usually derived from standard GK theory [119, 120]. As Palmer pointed out, the dynamics of the fluctuations is numerically hard to estimate from EMD simulations. Additionally, in the case of heat transport, energy density depends on the arbitrary partitioning of the total energy into atomic contributions [15, 119, 120].

In single component materials, the hydrodynamic equations governing the evolution of the densities are particularly simple [14]. In these systems, energy is uncoupled to other conserved quantities. Under these conditions, energy transport is a diffusive process [14]:

$$\frac{1}{\omega}\chi''_{nn}(\mathbf{k},\omega) = m\left.\frac{\partial n}{\partial p}\right|_T \left[\frac{c_v/c_p c^2 k^4 \Gamma_B}{(\omega^2 - c^2k^2)^2 + (\omega k^2\Gamma_B)^2} + \frac{(1-c_v/c_p)k^2 D_T}{\omega^2 + (k^2 D_T)^2} - \left(1 - \frac{c_v}{c_p}\right)\frac{(\omega^2 - c^2k^2)k^2 D_T}{(\omega^2 - c^2k^2)^2 + (\omega k^2\Gamma_B)^2}\right] \tag{4.1}$$

$$\frac{1}{\omega}\chi''_{qq}(\mathbf{k},\omega) = mnc_p T \frac{k^2 D_T}{\omega^2 + (k^2 D_T)^2} \tag{4.2}$$

where q is the heat density, n is the number density, $D_T = \frac{\kappa}{mnc_p}$, $\Gamma_B = D_l + D_T(c_p/c_v - 1)$ is the Brillouin peak and the other parameters are defined as in Eqs. (1.83), (1.85), (1.84). Leveraging the hydrodynamic functional form of the density-density correlation function, $\frac{1}{\omega}\chi''_{nn}$, Cheng and Frenkel proposed an accurate method to infer thermal conductivity of normal fluids from the wave-vector dependent width of the Rayleigh and Brillouin peaks [121],

$$\gamma_R(k) = D_T k^2 \tag{4.3}$$

$$\gamma_B(k) = \frac{1}{2}\left(D_l + D_T(c_p/c_v - 1)\right)k^2. \tag{4.4}$$

As a practical downside, the ratio between the intensities of the Brillouin and Rayleigh peak —the Landau Planczek (LP) ratio $R_{LP} = (c_p/c_v - 1)$— vanishes in incompressible fluids. For these systems, the analysis of the number density autocorrelation function does not provide any insight on the transport of heat [121]. Thanks to the work of Cheng, Frenkel and Palmer [14, 119–121], we develop a numerical technique for the extrapolation of thermal conductivity from energy density fluctuations in normal systems which overcomes some of the numerical drawbacks highlighted by Palmer and is applicable also to incompressible fluids and solids [122].

The Energy Density Structure Factor (EDSF), $\widetilde{C}(\mathbf{k}, \omega)$, is defined as

$$\frac{1}{\omega}\chi''_{ee}(\mathbf{k}, \omega) = \frac{1}{2k_B T}\widetilde{C}(\mathbf{k}, \omega), \tag{4.5}$$

$$\widetilde{C}(\mathbf{k}, \omega) = V\int_{-\infty}^{\infty} \langle \widehat{e}(\mathbf{k}, t)\widehat{e}(-\mathbf{k}, 0)\rangle\, e^{i\omega t} dt. \tag{4.6}$$

The continuity equation links in a very simple fashion the energy density, e, and current, j, correlations: $\widetilde{C}_{ee}(\mathbf{k}, \omega) = \frac{k^2}{\omega^2}\widetilde{C}_{jj}(\mathbf{k}, \omega)$. Thus, the extrapolation of heat conductivity from energy fluctuations is equivalent to the usual GK formula. In the hydrodynamic regime, the correlation function of the energy density is

$$\widetilde{C}(\mathbf{k}, \omega) = k_B T^2 \frac{k^2 D_T mnc_p}{\omega^2 + k^4 D_T^2}. \tag{4.7}$$

Therefore, heat conductivity can be extrapolated as:

$$\kappa = \lim_{\mathbf{k}\to 0} \kappa(\mathbf{k}), \tag{4.8}$$

$$\kappa(\mathbf{k}) = \frac{k_B T^2}{k^2}\frac{(mnc_p)^2}{\widetilde{C}(\mathbf{k}, \omega = 0)}. \tag{4.9}$$

On the theoretical side, as introduced by Palmer in 1994 [119, 120], this approach is simple and straightforward since it only assumes energy to be a diffusive process. On the numerical side, as discussed in Sec. 2.1, the zero-frequency limit in Eq. (4.9) inherits the computational limitations of time integrals of autocorrelation functions. We filter the EDSF via cepstral analysis [44] and then extrapolate the long-wavelength limit via Bayesian-regression analysis as detailed in Ref [75], Sec. 2.4 and Sec. 3.1 [122].

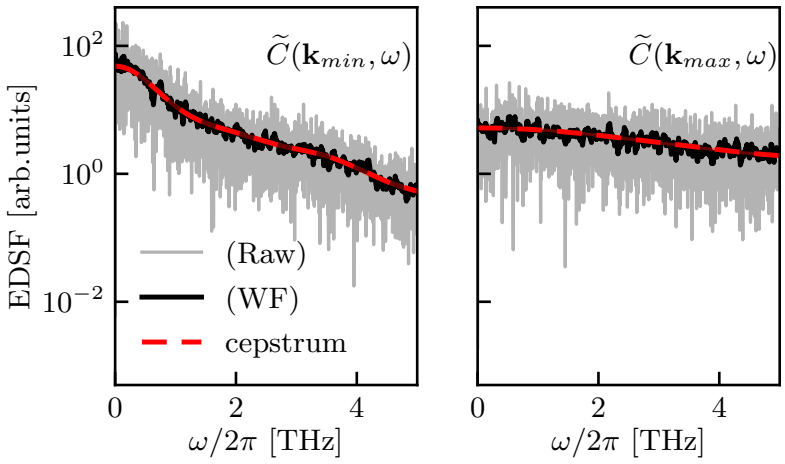

Figure 4.1: Upper panel: EDSF of liquid Argon computed at the **k**-vectors of minimum magnitude, $k_{min} = 1.77\,nm^{-1}$ [122]. Lower panel: same, but for the **k**-vectors, $k_{max} = 11.3\,nm^{-1}$ [122]. The WF line refers to a window filter of with 0.05 THz [122]. The shaded red area indicates the cepstral statistical uncertainty. Notice the logarithmic scale on the y-axis. Figure reproduced from Ref. [122].

As Cheng and Frenkel point out [121], even if energy currents and densities depend on the specific gauge choice, the long time integral of their long-wavelength autocorrelation function does not. For the current, the invariance is proved by the *gauge* invariance principle of transport coefficients [15]. One can arbitrarily add to the energy density the divergence of a bounded vector, $e'(\mathbf{r}) = e(\mathbf{r}) + \nabla \cdot \boldsymbol{p}(\mathbf{r})$. According to this prescription, the EDSF must change. Thanks to the Wiener-Khintchine theorem [51, 52], the $\omega = 0$ value of the transformed EDSF can be computed as

$$\widetilde{C}'(\mathbf{k},\omega=0) = \lim_{\tau\to\infty}\left[\frac{V}{2\tau}\left\langle\left|\int_0^\tau \widehat{e}(\mathbf{k},t)\,dt\right|^2\right\rangle + \left\langle\left|\int_0^\tau \mathbf{k}\cdot\widehat{\mathbf{p}}(\mathbf{k},t)\,dt\right|^2\right\rangle + 2\mathbf{k}\cdot \mathrm{Im}\left\langle\int_0^\tau \widehat{e}(\mathbf{k},t)\,dt\int_0^\tau \widehat{\mathbf{p}}(-\mathbf{k},t')\,dt'\right\rangle\right]. \tag{4.10}$$

In the limit $\mathbf{k}\to 0$ both the energy density's and $\mathbf{p}$'s Fourier transform are real. In the long-wavelength limit, Eq. (4.8) is unaffected by the gauge transformation.

We compute thermal conductivity of liquid Argon, liquid SPC/E water and, solid silica from the regression of energy density fluctuations [122]. As anticipated, we employ cepstral analysis to filter out the noise of the frequency/wave-vector dependent EDSF [122]. In EMD simulations, water is modeled with a sample of 216 molecules interacting through the SPC/E force field [89] at $T = 300$ K and $p = 0$ bar, and the trajectories are integrated using the velocity-Verlet algorithm as implemented in `LAMMPS` with a time-step

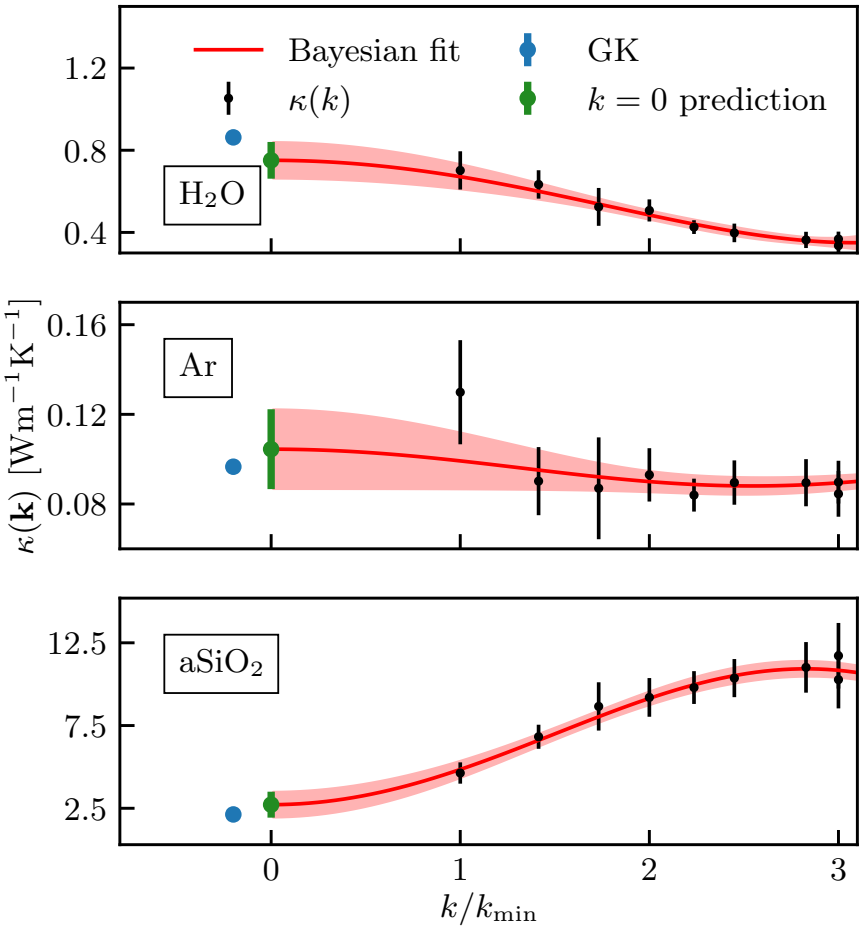

Figure 4.2: Upper panel: Wave-vector dependent heat conductivity, $\kappa(\mathbf{k})$, of SPC/E water at T = 300 K and 0 bar, estimated from the EMD simulation, $k_{min} = 3.40\,nm^{-1}$ [122]. Central panel: same for liquid Argon at T = 100 K and 0 bar, $k_{min} = 1.77\,nm^{-1}$ [122]. Lower panel: same for solid amorphous Silica at T = 500 K, $k_{min} = 3.04\,nm^{-1}$ [122]. The data are averages over equivalent wave-vectors. The red line indicates the Bayesian regression analysis with the associated estimated uncertainty. At $\mathbf{k} = 0$ we display the Bayesian extrapolation and GK estimate, on the same EMD trajectory, of heat conductivity [122]. Figure reproduced from Ref. [122].

of 0.25 fs [50, 122]. Argon is modeled with a sample of 864 atoms interacting through a Lennard-Jones potential [123] at T = 100 K and p = 0 bar, and time-step 1fs. Amorphous Silica is modeled with a sample of 648 atoms interacting via the BKS potential [124] at 500 K, and time-step 0.25 fs. The data are harvested from microcanonical simulations 1 ns, 2 ns and, 1 ns-long, of SPC/E water, liquid Argon and, solid Silica, respectively. The energy densities are sampled every 2.5, 2.5 and, 5 fs, respectively. [122].

In Fig. 4.1, the EDSF of liquid Argon displays the typical diffusion peak at zero-frequency, which is related to thermal diffusivity. The peak becomes sharper as we reduce the magnitude of the probed wave-vector, making the

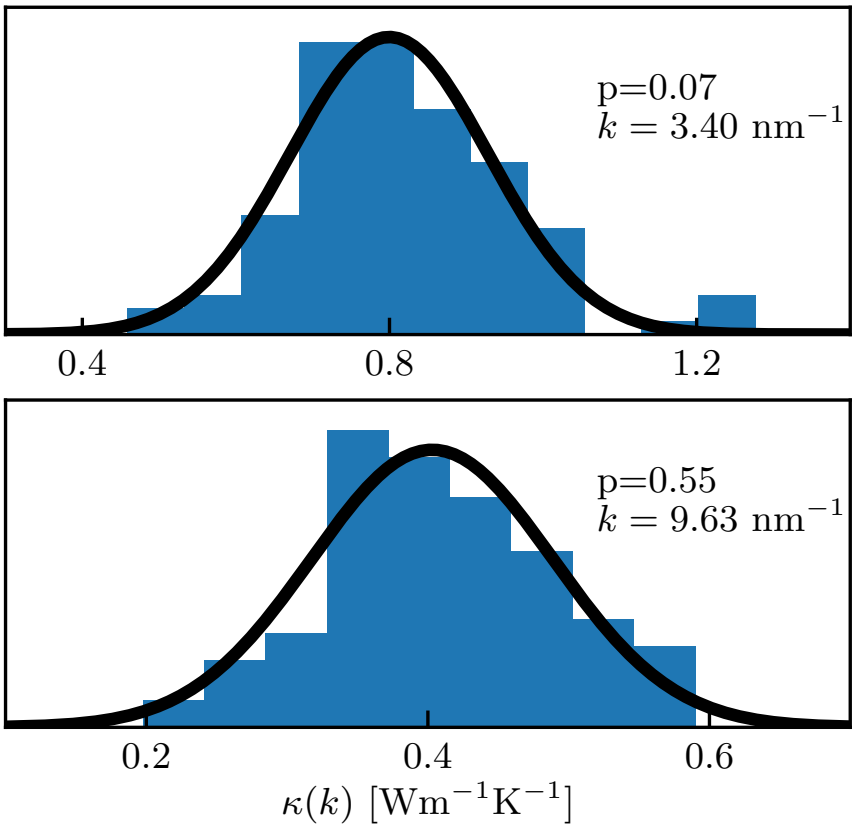

Figure 4.3: Distribution of the wave-vector dependent heat conductivity, $\kappa(\mathbf{k})$, of SPC/E water at T = 300K and 0 bar as evaluated from 100 200ps-long EMD trajectories and the p-value of the Shapiro-Wilk normality test [122]. The solid lines represent a Gaussian curve fitted on the distribution of heat conductivities at the wavevector presented [122]. Figure reproduced from Ref. [122].

cepstral procedure increasingly more challenging [122]. Once we have computed the zero-frequency limit via cepstral analysis, we are left with the non-trivial issue of estimating the long-wavelength limit [122].

As for the thermopolarization coefficient of liquid water, we rely on the Bayesian linear regression analysis described in Sec. 2.4 [122]. The wave-vector dependent thermal conductivity, Eq. 4.9, is approximated as a low order polynomial of magnitude of the $\mathbf{k}$-vector, Eq. (2.52),

$$\kappa(\mathbf{k}) \approx w_0 + w_1 \mathrm{k}^2 + w_2 \mathrm{k}^4 \cdots + w_M \mathrm{k}^{2M}.$$

In the Bayesian extrapolation, we optimize the hyperparameter α, which enters in the Gaussian prior of the parameters, Eq. (2.42). The optimization of the prior and the selection of the optimal model are discussed in Sec. 2.4. In Fig. 4.2 we present Eq. (4.9) and the Bayesian linear regressions [122], which are in good agreement with GK calculations estimated on the same trajectories. A crucial requirement for the long-wavelength regression described in

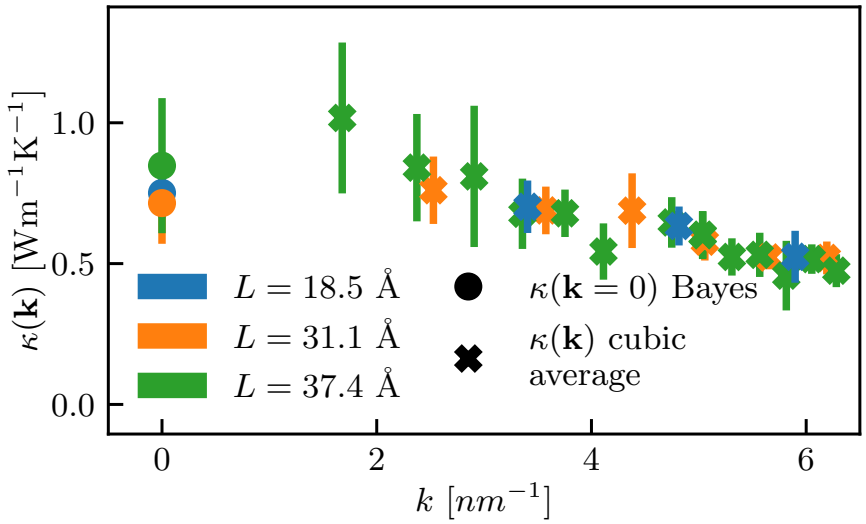

Figure 4.4: Wave-vector dependent heat conductivity, $\kappa(\mathbf{k})$, of SPC/E water at 300 K and 0 bar, as evaluated from the EMD simulations for three different linear box sizes, L: 18.5 Å, 31.1 Åand 37.4 Å; corresponding to 216, 1000 and 1728 water molecules respectively [122]. The data are averages over equivalent wave-vectors. At $\mathbf{k} = 0$ we report, as dots, the Bayesian extrapolations of heat conductivity at each size. Figure reproduced from Ref. [122].

Sec. 2.4 is the Gaussian distribution of the samples. According to Sec. 2.2, cepstral estimates of Eq. 4.9 are Gaussian variables. In order to investigate directly this aspect of the analysis, we evaluate thermal conductivity of SPC/E water from 100 200 ps-long trajectories [122]. Fig. 4.3 displays the histogram of the results of Eq. 4.9, which pass the Shapir-Wilk normality test with a level of significance of 0.05 [58, 122]. The long-wavelength extrapolation is stable as we increase the linear size of the simulation box [122]. In Fig. 4.4, we report the wave-vector dependent thermal conductivity and its long-wavelength Bayesian extrapolation for three increasing box sizes [122]. At finite wave-vectors, Eq. (4.9) estimates do not show significant finite size effects. Nevertheless, the error bar increases decreasing the magnitude of the probed $\mathbf{k}$-vector. The long-wavelength Bayesian extrapolation allows to restrict the analysis to small systems and short trajectory lengths [122].

Density-based methods designed for the computation of thermal conductivity, do not rely on the energy current, which for some classical force fields or quantum density functionals may be hard to express analytically or implement numerically. On a final note, we stress that the regression of energy density fluctuations is about one order of magnitude less efficient than the GK formula filtered via cepstral analysis, estimated on the same trajectory [122].

4.2 FOUNDATIONS OF QUASI-HARMONIC-GREEN-KUBO THEORY

Below the melting point, in solid insulators transport is governed by the vibrations of atoms oscillating around their equilibrium positions. Under the hypothesis that atomic vibrations are much smaller than interatomic distances, normal modes in the harmonic approximation, i. e. phonons, are the natural basis to describe the vibrational properties of the system. In this limit, one can take anharmonicity into account as a perturbation affecting the crystalline or amorphous solid [125, 126]. In strongly anharmonic systems, the quasi-harmonic approximation breaks down and other models of heat transport must be considered [127]. As in Refs.[125, 126], heat transport in solids is discussed according to the linear GK theory in the quasi-harmonic approximation. The potential energy, $\widehat{\mathcal{V}}$, of N atoms can be expanded around the equilibrium position, $\{\mathbf{R}_i^0\}$, as a functions of the displacements, $\{\widehat{\boldsymbol{u}}_i\}$,

$$\widehat{\mathcal{V}} = \mathcal{V}\left(\{\mathbf{R}^0\}\right) + \frac{1}{2}\sum_{i,j=1}^{N} \widehat{\boldsymbol{u}}_j^{\mathsf{T}} \cdot \left.\frac{\partial^2 \mathcal{V}}{\partial \boldsymbol{u}_i \partial \boldsymbol{u}_j}\right|_{\{\mathbf{R}^0\}} \cdot \widehat{\boldsymbol{u}}_i + \mathcal{O}\left(\widehat{\boldsymbol{u}}^3\right). \tag{4.11}$$

The second order derivatives of the potential with respect to the displacements, evaluated at the equilibrium positions, are usually referred to as second-order Interatomic Force Constants (IFC). Up to second-order corrections, we describe harmonic systems diagonalizing the dynamical matrix, i. e. the second-order IFC matrix rescaled by the masses of the atoms, $\mathbf{D}$,

$$D_{ij,\alpha\beta} = \frac{1}{\sqrt{M_i M_j}} \left.\frac{\partial^2 \mathcal{V}}{\partial \boldsymbol{u}_{i,\alpha} \partial \boldsymbol{u}_{j,\beta}}\right|_{\{\mathbf{R}^0\}} \tag{4.12}$$

$$\sum_{j\beta} D_{ij,\alpha\beta}\, \varepsilon_{j\mu,\beta} = \omega_\mu^2 \varepsilon_{i\mu,\alpha} \tag{4.13}$$

where ε_μ is the eigenvector with eigenvalue ω_μ^2. The normal modes and their amplitudes are naturally defined as

$$\widehat{\xi}_\mu = \sum_{i,\alpha} \varepsilon_{j\mu,\beta}\, \sqrt{M_i}\widehat{u}_{i,\alpha}, \quad \widehat{\pi}_\mu = \sum_{i,\alpha} \varepsilon_{j\mu,\beta}\, \sqrt{M_i}\dot{\widehat{u}}_{i,\alpha} \tag{4.14}$$

$$\widehat{a}_\mu = \sqrt{\frac{\omega_\mu}{2}}\widehat{\xi}_\mu + i\sqrt{\frac{1}{2\omega_\mu}}\widehat{\pi}_\mu. \tag{4.15}$$

In the basis of the normal amplitudes, the harmonic Hamiltonian is particularly simple:

$$\widehat{\mathcal{H}}_0 = \frac{1}{2}\sum_\mu \omega_\mu \left(\widehat{a}_\mu \widehat{a}_\mu^* + \widehat{a}_\mu^* \widehat{a}_\mu\right). \tag{4.16}$$

Thanks to the gauge invariance principle [19], the harmonic energy flux reads [125, 126]:

$$\widehat{J}_\alpha = \frac{1}{V}\sum_i R^0_{i,\alpha}\dot{\widehat{\epsilon}}_i \tag{4.17}$$

$$= \frac{1}{2V}\sum_{ij,\beta\gamma}\left(R^0_{i,\alpha}-R^0_{j,\alpha}\right)\sqrt{M_iM_j}D_{ij,\beta\gamma}\widehat{u}_{i,\beta}\dot{\widehat{u}}_{j,\gamma} \tag{4.18}$$

$$= \frac{i}{2V}\sum_{\mu\nu} v_{\mu\nu,\alpha}\omega_\mu\left(\widehat{a}^*_\mu+\widehat{a}_\nu\right)\left(\widehat{a}^*_\nu-\widehat{a}_\mu\right) \tag{4.19}$$

where the minimum image convention is understood computing the position differences and $v_{\mu\nu,\alpha}$ is the first moment of the dynamical matrix, i.e. the velocity matrix,

$$v_{\mu\nu,\alpha} = \frac{1}{2\sqrt{\omega_\mu\omega_\nu}}\sum_{ij,\beta\gamma}\left(R^0_{i,\alpha}-R^0_{j,\alpha}\right)\sqrt{M_iM_j}D_{ij,\beta\gamma}\varepsilon_{i\nu,\beta}\varepsilon_{j\mu,\gamma}. \tag{4.20}$$

The anharmonic perturbation to the harmonic Hamiltonian reads

$$\widehat{\mathcal{H}}''' = \frac{1}{6}\sum_{ijk,\alpha\beta\gamma}\frac{\partial^3\mathcal{V}}{\partial u_{i,\alpha}\partial u_{j,\beta}\partial u_{k,\gamma}}\widehat{u}_{i,\alpha}\widehat{u}_{j,\beta}\widehat{u}_{k,\gamma}. \tag{4.21}$$

We treat Eq. (4.21) as an external perturbation affecting the harmonic Hamiltonian. The anharmonic disturbance can be decomposed in the the basis of the normal amplitudes,

$$\widehat{\mathcal{H}}''' = \sum_{\mu\nu\lambda}\Phi_{\mu\nu\lambda}\frac{1}{\sqrt{\omega_\mu\omega_\nu\omega_\lambda}}\left(\widehat{a}_\mu+\widehat{a}^*_\mu\right)\left(\widehat{a}_\nu+\widehat{a}^*_\nu\right)\left(\widehat{a}_\lambda+\widehat{a}^*_\lambda\right) \tag{4.22}$$

where $\Phi_{\mu\nu\lambda}$ is the third-order IFC tensor computed with respect to the normal modes coordinates. The quantum analogue of the previous relations is obtained mapping the amplitudes into the bosonic creator/annihilation operators: $\widehat{a}_\mu;\ \widehat{a}^*_\mu \mapsto \sqrt{\hbar}\widehat{a}_\mu;\ \sqrt{\hbar}\widehat{a}^\dagger_\mu$. According to GK linear response theory, thermal conductivity is

$$\kappa_{\alpha\beta} = \begin{cases} \frac{1}{TV}\int_0^{\frac{1}{k_BT}} ds \int_0^\infty dt \left\langle \widehat{J}_\alpha(t-i\hbar s)\widehat{J}_\beta(0)\right\rangle; & \text{quantum} \\ \frac{1}{k_BT^2V}\int_0^\infty dt \left\langle \widehat{J}_\alpha(t)\widehat{J}_\beta(0)\right\rangle; & \text{classical} \end{cases}. \tag{4.23}$$

Since the harmonic energy flux is a quadratic in the amplitudes, or creator-annihilation operators, the average in Eq. (4.23) is a 4-body correlation function. We factorize the 4-body correlation function via the Wick theorem, or classically by the Isserlis theorem, as products of 2-body correlations [125],

$$\left\langle \widehat{a}^\dagger_\mu(t)\widehat{a}_\nu(t)\widehat{a}^\dagger_\sigma\widehat{a}_\lambda\right\rangle \approx \left\langle \widehat{a}^\dagger_\mu\widehat{a}_\nu\right\rangle\left\langle \widehat{a}^\dagger_\sigma\widehat{a}_\lambda\right\rangle\delta_{\mu\nu}\delta_{\sigma\lambda} + \left\langle \widehat{a}^\dagger_\mu(t)\widehat{a}_\lambda\right\rangle\left\langle \widehat{a}_\nu(t)\widehat{a}^\dagger_\sigma\right\rangle\delta_{\mu\lambda}\delta_{\nu\sigma}. \tag{4.24}$$

As a further approximation, we model the decay of the phonons [125, 126]. In order to do so, we assume a Markovian decay, neglecting memory effects [125, 126]:

$$\left\langle \widehat{a}_\mu^\dagger \widehat{a}_\mu(t) \right\rangle \approx n_\mu e^{-i\omega_\mu t - |t|\gamma_\mu}. \tag{4.25}$$

The linewidth, γ_μ, is the decay rate of the mode μ due to the anharmonic interactions and, in the quasi-harmonic limit, $\gamma_\mu/\omega_\mu \ll 1$. Under these hypotheses, thermal conductivity is [125, 126]

$$\kappa_{\alpha\beta} = \frac{1}{V}\sum_{\mu\nu} \frac{n_\mu(n_\nu+1)+n_\nu(n_\mu+1)}{2k_B T^2}\frac{(\omega_\mu+\omega_\nu)^2}{4} v_{\mu\nu,\alpha} v^*_{\mu\nu,\beta}\tau_{\mu\nu} \tag{4.26}$$

$$\tau_{\mu\nu} \equiv \frac{\gamma_\mu+\gamma_\nu}{(\omega_\mu-\omega_\nu)^2+(\gamma_\mu+\gamma_\nu)^2} \tag{4.27}$$

$$C_{\mu\nu} \equiv \frac{n_\mu(n_\nu+1)+n_\nu(n_\mu+1)}{2k_B T^2}\frac{(\omega_\mu+\omega_\nu)^2}{4}. \tag{4.28}$$

Thanks to the Fermi Golden Rule (FGR), the anharmonic linewidths read [125, 126]

$$\gamma_\mu = \hbar\pi \sum_{\mu\lambda} \left|\frac{\Phi_{\mu\nu\lambda}}{\sqrt{\omega_\mu\omega_\nu\omega_\lambda}}\right|^2 \left(\frac{1}{2}(n_\nu+n_\lambda+1)\,\delta(\omega_\mu-\omega_\nu-\omega_\lambda) \right. \\ \left. (n_\nu-n_\lambda)\,\delta(\omega_\mu+\omega_\nu-\omega_\lambda)\right). \tag{4.29}$$

Thermal conductivity in the Quasi Harmonic Green Kubo (QHGK) theory is [125, 126, 128]:

$$\kappa_{\alpha\beta} = \frac{1}{V}\sum_{\mu\nu} C_{\mu\nu} v_{\mu\nu,\alpha} v^*_{\mu\nu,\beta}\tau_{\mu\nu}. \tag{4.30}$$

QHGK theory addresses at once both crystalline and disordered systems, connecting the Allen Feldman (AF) theory of amorphous solids [129–133] and the linearized Boltzmann Transport Equation (BTE) in the Relaxation Time Approximation (RTA), designed for crystals [128, 134–136]. The limitation of BTE have been overcome in the recently proposed Wigner Boltzmann Transport Equation (WBTE) formalism [137–141].

In amorphous systems, harmonic normal modes are not diagonal in the plain waves basis, but at low frequencies. Indeed, in such solids, low-energy vibrations are responsible for sound propagation. The AF model is a purely harmonic theory of heat propagation in glasses. In the limit of vanishing anharmonic linewidths, thermal conductivity is

$$\kappa_{\alpha\beta} = \frac{1}{V}\sum_{\mu} C_\mu \mathcal{D}_\mu, \tag{4.31}$$

where C_μ is the specific heat of the mode μ, $C_\mu = \hbar\omega_\mu \frac{\partial n_\mu}{\partial T}$, and $\mathcal{D}_\mu$ is the diffusivity of the mode, defined as,

$$\mathcal{D}_\mu = \pi \sum_\nu v_{\mu\nu,\alpha} v^*_{\mu\nu,\beta} \delta(\omega_\mu - \omega_\nu) \tag{4.32}$$

$$= \lim_{\eta\to 0\, V\to\infty} \sum_\nu v_{\mu\nu,\alpha} v^*_{\mu\nu,\beta} \frac{\eta}{(\omega_\nu - \omega_\mu)^2 - \eta^2}, \tag{4.33}$$

η being the smearing parameter to be made arbitrary small while increasing the size of the system. One can decompose thermal conductivity into modes contributions. Allen and Feldman provided a qualitative classification of modes in glasses as *propagons*, *diffusons* and *locons*[129, 131, 133]. The distinction is based on the Vibrational Dynamical Structure Factor (VDSF), $S_b(\mathbf{k}, \omega)$, which for harmonic systems reads

$$S^\circ_b(\mathbf{k}, \omega) = \sum_\mu \delta(\omega - \omega_\mu) |\langle \mu | \mathbf{k}, b \rangle| \tag{4.34}$$

$$\langle \mu | \mathbf{k}, b \rangle = \frac{1}{\sqrt{N}} \sum_{i,\alpha} \zeta_{b,\alpha}(\mathbf{k}) \varepsilon^*_{i\mu,\alpha} e^{i\mathbf{k}\cdot\mathbf{R}_i}, \tag{4.35}$$

where b is the brach index, $\mathbf{R}_i$ is the position of the i-th atom, $\mathbf{k}$ is the wavevector compatible with PBCs and $\zeta_b(\mathbf{k})$ is the polarization unit vector. The harmonic VDSF measures the overlap between of the modes with frequency ω_μ and the plain wave. Anharmonic effects can be included in the RTA approximation as:

$$S_b(\omega, \mathbf{k}) = \frac{1}{\pi} \sum_\mu \frac{\gamma_\mu}{\gamma_\mu^2 + (\omega - \omega_\mu)^2} |\langle \mu | \mathbf{k}, b \rangle|^2. \tag{4.36}$$

In the low-frequency regime, the VDSF can be fitted as a narrow Lorentzian centered in ω_μ,

$$S_b(\omega, \mathbf{k}) \approx \frac{\alpha_b(\mathbf{k})}{\pi} \frac{\Gamma_b(\mathbf{k})}{(\omega - c_b k)^2 + \Gamma_b(\mathbf{k})^2}. \tag{4.37}$$

The low-energy excitations are dubbed propagons, they are responsible for sound propagation, have a linear dispersion and a population density quadratic in frequency. Increasing the energy, we encounter an intermediate region where the $\mathbf{k}$-vector is almost a good quantum number, linearity breaks down but the linewidth is still smaller than the probed frequency. Finally, in the high-energy sector, locons are present, which are localized states fully spread in the Fourier space [131, 133, 142].

4.3 THE ROLE OF ANHARMONICITY IN HEAT TRANSPORT IN GLASSES

Because of the intrinsic disordered structure of glasses, in order to to sample the low-energy vibrations large size models are required. To overcome this computational drawback, hydrodynamic extrapolations of the propagons'

contribution to thermal conductivity have been developed [142]. Finite models of glasses implicitly introduce a infrared cutoff below the minimal available frequency $\omega_{\min} = c\frac{2\pi}{L}$, where c is the speed of sound and $\frac{2\pi}{L}$ is the minimum magnitude of the wave-vectors compatible with PBC.

As soon as the AF harmonic theory of glasses was proposed, it was realized that, including only Rayleigh (ω^4) scattering, a continuous Debye approximation of hydrodynamic vibrations results in the infrared divergence of bulk thermal conductivity at all temperature, irrespective of configurational disorder [129–131, 133, 142–145]. Extensive numerical simulations are required to support the hypothesis of Rayleigh scattering, but direct numerical diagonalization of the dynamical matrix for models with more than 20000 atoms is practically unfeasible. Therefore, the infrared divergence is very challenging to observe in numerical simulations.

By iteratively estimating the harmonic VDSF, a Lanczos-Haydock procedure circumvents the computational impediment due to large sizes. The harmonic VDSF associated with propagons can be approximated as a Lorentzian:

$$S_b^\circ(\omega, \mathbf{k}) \approx \frac{\alpha_b^\circ(\mathbf{k})}{\pi} \frac{\Gamma_b^\circ(\mathbf{k})}{(\omega - c_b k)^2 + \Gamma_b^\circ(\mathbf{k})^2}, \tag{4.38}$$

whose parameters are fitted from the raw data obtained from either direct numerical diagonalization or Haydock's method.

In the bulk limit, the number of atoms is extended to infinity, allowing the sum over the modes to be converted into a frequency integral. The contribution of propagons to thermal conductivity is

$$\kappa_P = \sum_b \frac{c_b^2}{3} \int_0^{\omega_P} C(\omega)\rho_b(\omega) \frac{1}{2\Gamma_b(\omega/c_b)} d\omega. \tag{4.39}$$

The propagons' limit frequency, ω_P, which signals the end of the low-frequency regime and introduces diffusons, is a tunable parameter, to be selected carefully via convergence analysis [142, 145–147]. The continuous integral, reflecting the bulk limit of propagons' contribution, is usually referred to as hydrodynamic extrapolation [142]. Obviously, the sampling of diffusons (and locons) is not affected by the infrared cutoff, thus their contribution, κ_D, can be estimated from medium/small size glasses. In the anharmonic approximation, grounded in hydrodynamic considerations, the low-frequency functional form of the sound attenuation linewidth is [142, 148]

$$\Gamma_b(\omega/c_b) \approx A_b\omega^2 + B_b\omega^4. \tag{4.40}$$

In the harmonic limit, $A_b = 0$, only Rayleigh scattering is present. The anharmonic lifetimes are estimated from the FGR on small systems and then fitted at low frequencies. By briefly inspecting Eq. (4.39), noticing that the propagons' density of states is quadratic and that the mode heat capacity $C(\omega) \propto \omega$, the pure harmonic ω^4 disorder-induced Rayleigh scattering would inevitably cause thermal conductivity to diverge. Random media theory also predicts the occurrence of Rayleigh, ω^4, sound attenuation in harmonic systems [149, 150].

Being prototypical examples of glasses, widely addressed in the scientific literature both computationally and experimentally, amorphous silica ($aSiO_2$), silicon (aSi) and, silicon carbide (aSiC) are considered in the present study [129, 130, 133, 145, 151–163]. We present extensive numerical simulations validating the presence of Rayleigh scattering in harmonic glasses [145]. Haydock's methods can address the harmonic VDSF of large models of disordered systems, with almost 140000 atoms.

The diagonal elements of the harmonic vibrational Green's function are of the form:

$$\lim_{\eta\to 0} \operatorname{Im} \langle \mathbf{k}, b| \left((\omega + i\eta)^2 - \mathbf{D}\right)^{-1} |\mathbf{k}, b\rangle = \frac{\pi}{2|\omega|} \left[S_b^0(\omega, \mathbf{k}) + S_b^0(-\omega, \mathbf{k})\right]. \quad (4.41)$$

In Haydock's approach, the harmonic VDSF is computed as a continued fraction:

$$\frac{\pi}{2|\omega|} \left[S_b^0(\omega, \mathbf{Q}) + S_b^0(-\omega, \mathbf{Q})\right] = \lim_{\eta\to 0} \operatorname{Im} \cfrac{1}{(\omega + i\eta)^2 - a_0 - \cfrac{b_1^2}{(\omega + i\epsilon)^2 - a_1 - \cfrac{b_2^2}{\ddots}}}, \quad (4.42)$$

whose coefficients $\{a_0, a_1, \dots\}$ and $\{b_1, b_2, \dots\}$ are estimated from the iterative Lanczos chain:

$$\begin{aligned} &|\xi_{-1}\rangle = 0, \\ &|\xi_0\rangle = |\mathbf{Q}, b\rangle \\ &b_n |\xi_n\rangle = (\mathbf{K} - a_{n-1}) |\xi_{n-1}\rangle - b_{n-1} |\xi_{n-2}\rangle \\ &a_n = \langle \xi_n | \mathbf{D} | \xi_n \rangle \\ &b_n = \langle \xi_n | \mathbf{D} | \xi_{n-1} \rangle . \end{aligned} \quad (4.43)$$

Haydock's iterative routine significantly alleviates the computational burden of harmonic VDSFs, which, in direct diagonalization algorithms, requires $\mathcal{O}((3N)^3)$ operations, to be compared with the $\mathcal{O}((3N)^2)$ complexity of a single Lanczos step. Typical simulations converge after $k_L = 600$ Lanczos steps, significantly reducing the numerical overload to $\mathcal{O}(k_L(3N)^2))$, k_L being much smaller than N. Since $\mathbf{D}$ is sparse, the number of operations needed by Haydock's routine can be further simplified to a complexity $\mathcal{O}(k_L 3N)$. Haydock's procedure is commonly used for diagonalization purposes but is notoriously unstable in the estimation of eigenvectors [164, 165]. The iterative algorithm is carefully benchmarked, computing the VDSF of a model of aSiC with 13824 atoms, against direct diagonalization. In Fig. 4.5 we display the sound damping coefficients, showing excellent agreement between the two methods [145].

The numerical simulation in support of the Rayleigh scattering hypothesis in harmonic systems are estimated from 13824, 97336 and, 139968-atoms models of aSi, aSiC and, $aSiO_2$, respectively; and displayed in Fig. 4.6. aSi is highly disordered and a model of roughly 10000 atoms suffices to show Rayleigh scattering. On the contrary, $aSiO_2$ hinders the truly harmonic ω^4 trend, up

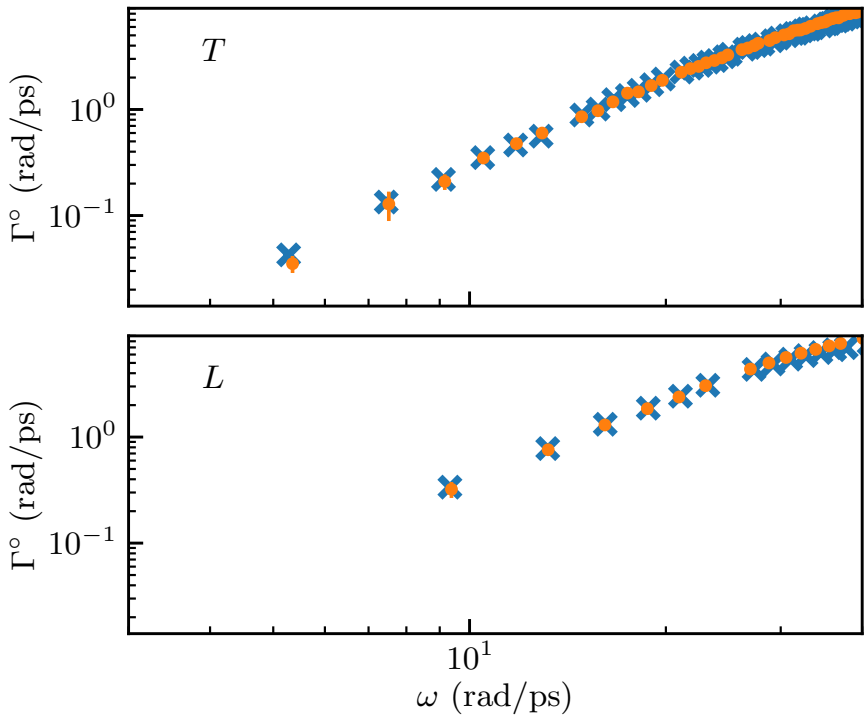

Figure 4.5: Harmonic sound damping coefficients of aSiC obtained by fitting the harmonic vibrational dynamical structure factor computed via Haydock's method (orange dots) and via direct diagonalization (cyan crosses). Both results are averaged over 4 samples of 13,824 atoms, the error bars represent standard deviations. Upper panel, transverse modes; lower panel, longitudinal modes. Figure reproduced from Ref. [145]

to 140000 atoms models. aSiC is an intermediate system, which nevertheless requires at least 30000 atoms to reveal the ω^4 scaling. aSi interatomic force field is modeled with the Tersoff force field [166], aSiC and $aSiO_2$ with the Vashishta potential [167, 168]. Our results are averaged over 4, 4 and, 10 independent samples of $aSiO_2$, aSiC and, aSi, respectively. Further details on the numerical simulations and the quenching procedure for generation of the glasses are reported in Ref. [145]. Fig. 4.6 shows the harmonic sound dumping coefficients, which displays a crossover $\omega^4 \to \omega^2$. Taking into account the crossover in harmonic systems, one has [153]

$$\Gamma_b^\circ(\omega) = C_b \omega^2 [1 + (\omega_{XO}^b/\omega)^{2\delta}]^{-1/\delta}, \tag{4.44}$$

where C_b is a constant, ω_{XO}^b is the polarization-dependent crossover angular frequency, and $\delta = 1.5$ determines the sharpness of the transition.

The AF thermal conductivity, Eqs. (4.39) (4.31) (4.32) (4.33), can be studied as a function of the smearing η and of the temperature T:

$$\kappa = \frac{1}{3V} \sum_{\mu\nu} C_\mu |v_{\mu\nu}|^2 \frac{\eta}{(\omega_\mu - \omega_\nu)^2 + \eta^2}. \tag{4.45}$$

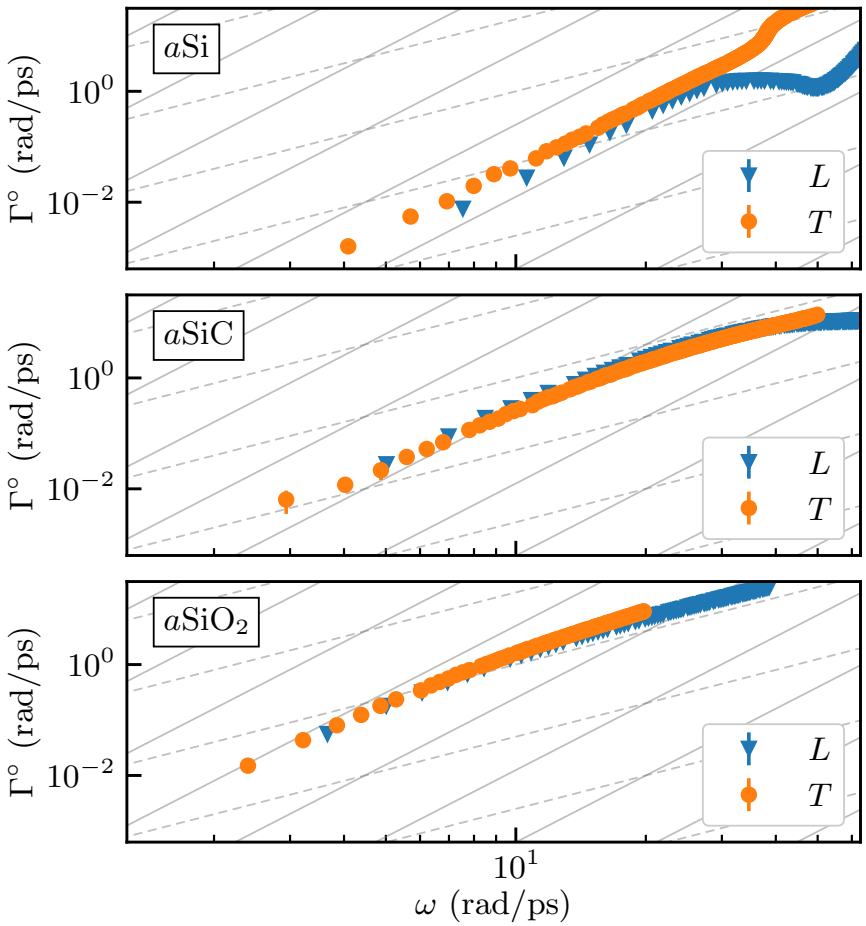

Figure 4.6: Sound damping coefficients in the harmonic approximation for aSi, aSiC, and $aSiO_2$. The estimate is obtained by fitting the harmonic vibrational dynamical structure factor, and expressing the linewidth as a function of ω. The sizes of the models are respectively 13824, 97336 and 139968 atoms. The estimated errors are smaller than the size of the markers. The dashed and continuous gray lines represent respectively the ω^2 and ω^4 scaling. Note the logarithmic scale on both axes. Figure reproduced from Ref. [145]

As displayed in Fig. 4.7, in finite systems, as the AF linewidth vanishes, the Lorentzian window encompasses less and less modes, causing thermal conductivity to be artificially reduced. Keeping the smearing fixed and increasing the number of atoms, we achieve converge in size, reaching the hydrodynamic limit described in Eq. 4.39 [145]. We fit the harmonic sound dumping coefficients, presented in Fig. 4.6, according to Eq. (4.44) [145]. In the case of aSiC and $aSiO_2$ the crossover angular frequency is $\omega_{XO} = 12\,\mathrm{rad\,ps^{-1}}$ and $6\,\mathrm{rad\,ps^{-1}}$. Fitting the harmonic sound dumping coefficients, one infers the propagons' contribution to thermal conductivity at low frequency as defined in Eq. (4.39). In Fig. 4.7 we show that in the bulk and $\eta \to 0$ limit, AF thermal conductivity diverges at all temperatures, due to the infrared contribution of propagons described in Eq. (4.39).

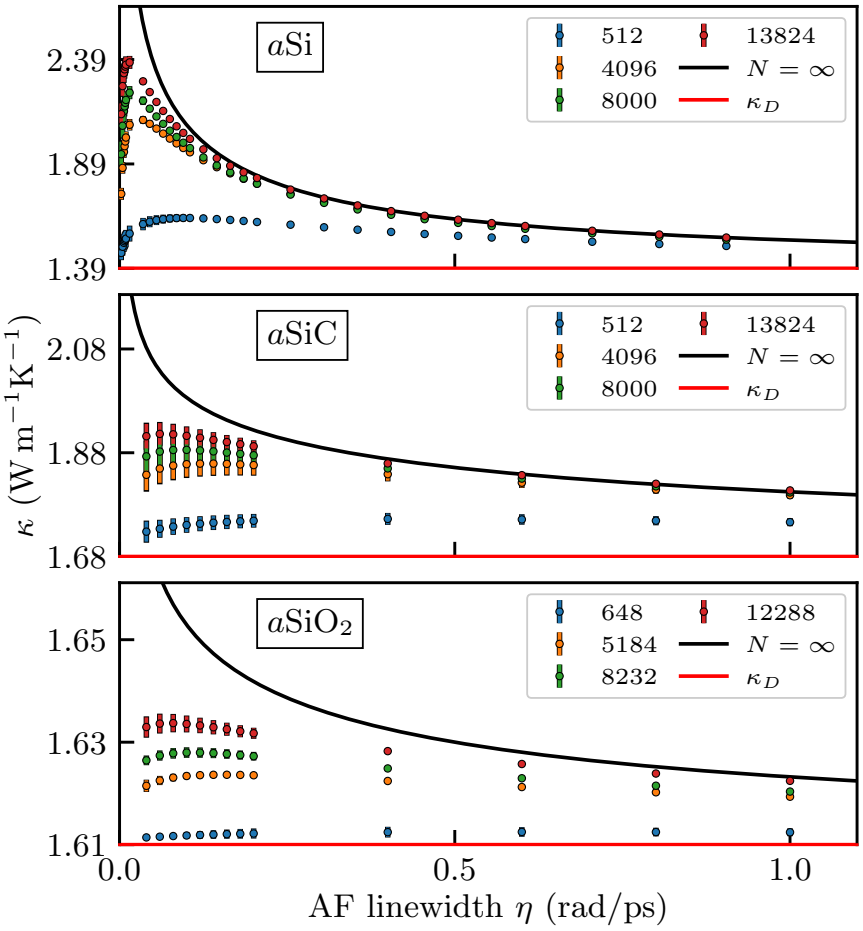

Figure 4.7: AF thermal conductivity of aSi, aSiC, and aSiO$_2$ samples with different number of atoms (N), as a function of the smearing, η. Calculations are made at a temperature of 500 K. The black, solid, line is the sum of the infinite-size analytical result for propagon and the small-size term due to diffusons/locons. The red solid line is the diffusons' contribution. The cutoff frequency, ω_P, for propagons is set to $\omega_P/2\pi = 3$ THz, $\omega_P/2\pi = 3$ THz, and $\omega_P/2\pi = 1.2$ THz, for the three materials, respectively. The red solid line indicates the diffusons' contribution to thermal conductivity, κ_D. Figure reproduced from Ref. [145]

In the harmonic approximation, the constant smearing factor employed in the AF calculations can be associated to a boundary scattering contribution present in thin-films glasses [145]. The finite size of the samples sets an upper-bound to the average mean free path of the modes [145]. One can map formally the AF smearing to the thickness of the thin-films [145]. In typical experiments, the glass samples are 1 nm-80 µm thick. To compare the numerical results to experiments, in the QHGK approximation we include anharmonic linewidths [145]. We compute the QHGK thermal conductivity in small systems and then fit the anharmonic linewidths at low frequencies [142, 145]. According to the Mattihessen sum rule, the results of the fit are combined with the

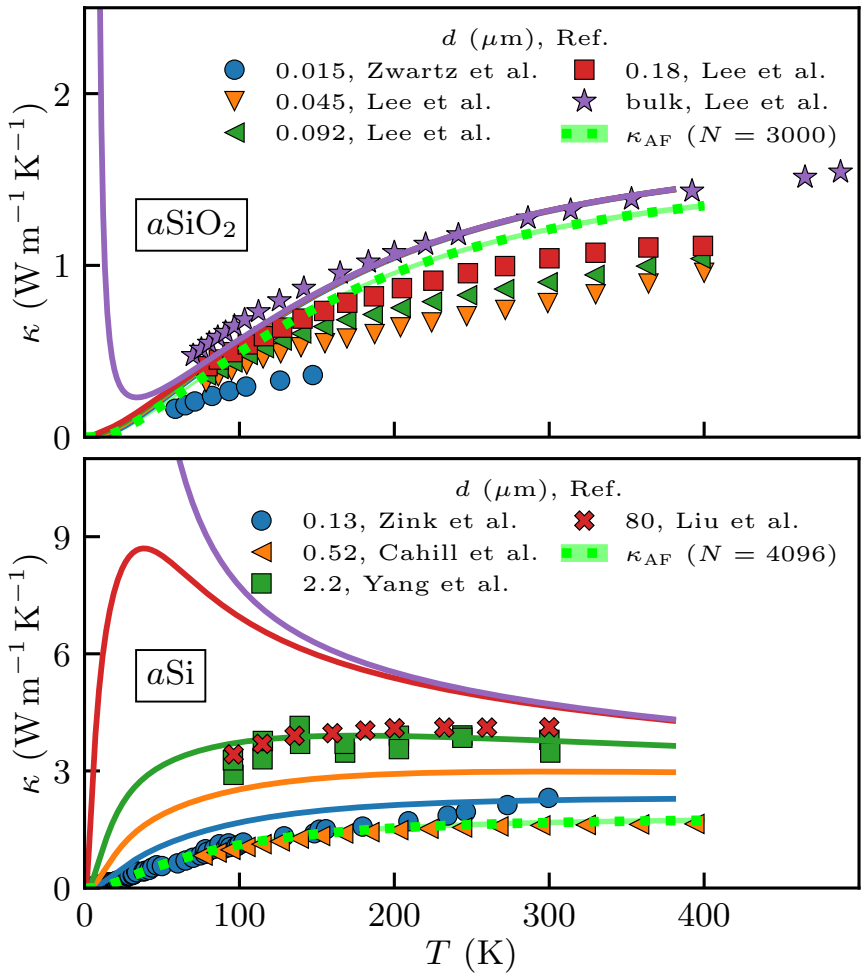

Figure 4.8: Thermal conductivity of $a\mathrm{SiO_2}$ and aSi as functions of temperature. Markers are experimental data taken from Refs. [163, 169] ($a\mathrm{SiO_2}$) and Refs. [157, 160–162] (aSi). The solid lines, color-matched to the markers, represent hydrodynamic QHGK results for films of the same thickness. Light-green, dotted lines are AF results on a finite system, and obtained with a smearing parameter η set to the average angular frequency spacing. Figure reproduced from Ref. [145]

harmonic ones, including also boundary scattering effects of thin films [145, 170].

The QHGK calculation, with the addition of the hydrodynamic extrapolation and boundary scattering, are presented in Fig. 4.8 and compared with experiments on thin-films, in a wide range of thicknesses and temperatures. In the bulk limit at finite temperature, the anharmonic linewidths, which scale as ω^2, prevent bulk thermal conductivity to dramatically diverge. At zero temperature and infinite size, neglecting quantum tunneling, the system becomes harmonic [142, 145]. In thin-films the boundary scattering is independent of temperature and, as T approaches zero, the occupation of the modes is dras-

tically reduced, maintaining a constant lifetime. Therefore, in thin-films, even in the quasi-harmonic regime, thermal conductivity vanishes at zero temperature.

In conclusion, this work proves that in the harmonic approximation, bulk thermal conductivity diverges at all temperatures due to the unlimited contribution of propagons [145]. The infrared catastrophe was predicted by the same Allen and Feldman in their seminal papers [129, 131], but, due to the limited computational resources, was never addressed in actual numerical simulations. We have discussed in details the numerical algorithm which allows us to perform VDSF calculations on models of glasses with almost 140000 atoms [145]. The effectiveness of previous AF calculations is rationalized in terms of boundary scattering and comparison to experiments on thin-films, whose thickness can be directly mapped into a finite AF smearing. The hydrodynamic extrapolation emphasizes the importance of accurate quasi-harmonic linewidths, which make bulk thermal conductivity, divergent in the harmonic case, finite at non zero temperatures [145].

CONCLUSIONS

Although transport coefficients are crucial for technological applications, their computation using EMD simulations requires long trajectories. In this Thesis, thermal transport and thermoelectric effects have been studied in both classical fluids and solids.

On one hand, in ionic conductors a temperature gradient generates a steady electric current in closed circuit conditions. On the other hand, in insulators charge transport is forbidden. Nevertheless, in polar fluids the molecules can rearrange to induce a macroscopic polarization across the system. Standard transport theory links the Seebeck coefficient to the ratio of an off-diagonal and a diagonal element of the Onsager matrix. In the absence of charge transport, this formulation results in an indeterminate form $\frac{0}{0}$. I developed a static theoretical framework to investigate thermopolarization in polar fluids from EMD simulations. In insulators, the dielectric response to a temperature inhomogeneity is a quasi-equilibrium phenomenon. Under these conditions, when perturbed, dielectric degrees of freedom relax in a microscopic time scale. Thermopolarization, being the response to a homogeneous perturbation —incompatible with PBC— is particularly hard to access. Thanks to static linear response theory, the thermopolarization coefficient is estimated as the long-wavelength limit of a suitable periodic static susceptibility. I employed Bayesian regression analysis to evaluate the long-wavelength extrapolation and its statistical uncertainty.

In ionic conductors, the Seebeck effect is a fundamental thermoelectric phenomenon. Since the Seebeck coefficient is derived from the Onsager matrix, we designed a Bayesian regression protocol to filter the spectra of both diagonal and off-diagonal Onsager coefficients. In a recently developed all-encompassing Bayesian regression framework, a single statistical model is able to address the entire Onsager matrix. The spectral methods presented in this Thesis overcome some of the numerical drawbacks that hindered the computation of transport coefficients via EMD methods.

The final chapter of this Thesis addresses heat transport. In solids, in the quasi-harmonic approximation, QHGK theory treats both crystalline and disordered systems in a unified formalism. Investigating exceptionally large harmonic models of glasses, we numerically verified the presence of disorder-induced Rayleigh scattering in the propagon regime. The existence of Rayleigh scattering at low frequencies has far-reaching consequences. In harmonic systems, it predicts a divergent bulk thermal conductivity at all temperatures, critically emphasizing the importance of anharmonic effects.

APPENDIX

A

KUBO'S AND MARTIN'S DISCUSSION ON TRANSPORT THEORY

THE APPENDIX reports results and discussions presented in Ref. [11] and Ref. [13]. In 1965, Martin proposed a relationship between the response to slowly varying external forces and to the deviations of internal observables [13]. In 1966 Kubo discussed a Langevin approach to tackle the same problem [11]. The method is discussed in details in Ref. [11] and here I review some of the results.

In an interacting system where the Hamiltonian is perturbed by $\widehat{\mathcal{H}}' \propto \widehat{b}(-\mathbf{k})v_b^{ext}(\mathbf{k})$, the continuity equation of the density b is

$$\widehat{\dot{b}}(\mathbf{k},t) = -i\mathbf{k}\cdot\widehat{\boldsymbol{j}}_b^0(\mathbf{k},t) - i\mathbf{k}\cdot\widehat{\boldsymbol{j}}'_b(\mathbf{k},t), \tag{A.1}$$

Eq. A.1 is equivalent to a Langevin equation such that

$$\begin{aligned} -i\mathbf{k}\cdot\widehat{\boldsymbol{j}}_b^0 &= -\int_0^t \gamma(\mathbf{k},t-t')\widehat{b}(\mathbf{k},t')\,dt' \\ \widehat{\boldsymbol{j}}_b(\mathbf{k},t) &= \widehat{\boldsymbol{j}}_b^0(\mathbf{k},t) + \widehat{\boldsymbol{j}}'_b(\mathbf{k},t), \end{aligned} \tag{A.2}$$

where $\boldsymbol{j}_b^0(\mathbf{k},t)$ and $\boldsymbol{j}'_b(\mathbf{k},t)$ are the systematic and random currents, respectively [11].

The response of the current to the gradient of the external potential,

$$\left\langle \widehat{\boldsymbol{j}}_b(\mathbf{k},t)\right\rangle = -\operatorname{Re}\mu^{ext}(\mathbf{k},\omega)i\mathbf{k}v_b^{ext}(\mathbf{k})e^{i\omega t}. \tag{A.3}$$

is computed from GK theory as [11],

$$\mu^{ext}(\mathbf{k},\omega) = V\beta\int_0^\infty \left\langle \widehat{\boldsymbol{j}}_b(-\mathbf{k})\widehat{\boldsymbol{j}}_b(\mathbf{k},t)\right\rangle_0 e^{i\omega t}\,dt\,. \tag{A.4}$$

$\gamma(\mathbf{k},t)$ in Eq. (A.2), can be evaluated from the fluctuations of the random current $\boldsymbol{j}'_b(\mathbf{k},t)$ as [11]

$$\gamma(\mathbf{k},\omega) = \frac{1}{\chi_{bb}(\mathbf{k},0)}\mathbf{k}\cdot V\beta\int_0^\infty \left\langle \widehat{\boldsymbol{j}}'_b(-\mathbf{k})\widehat{\boldsymbol{j}}'_b(\mathbf{k},t)\right\rangle_0 e^{i\omega t}\,dt\cdot\mathbf{k}, \tag{A.5}$$

$\chi_{bb}(\mathbf{k},\omega)$ being the Fourier transform of the response function defined in Ch. 1. Thanks to Eq. (A.5), one can define the screened mobility, μ^* [11]:

$$\gamma(\mathbf{k},\omega) \equiv \frac{\mathbf{k}\cdot\mu^*(\mathbf{k},\omega)\cdot\mathbf{k}}{\chi_{bb}(\mathbf{k},0)}. \tag{A.6}$$

Ref. [11] proves that the relation between μ^{ext} and μ^* is [11, 13]:

$$\mu^*(\mathbf{k},\omega) = \frac{\mu^{ext}(\mathbf{k},\omega)}{1-\frac{\mathbf{k}\cdot\mu^{ext}(\mathbf{k},\omega)\cdot\mathbf{k}}{i\omega\chi_{bb}(\mathbf{k},0)}}. \tag{A.7}$$

One can define the *shielding factor*, $\epsilon^*(\mathbf{k},\omega)$,

$$\mu^{\text{ext}}(\mathbf{k},\omega) = \frac{\mu^*(\mathbf{k},\omega)}{\epsilon^*(\mathbf{k},\omega)}. \tag{A.8}$$

$\epsilon^*(\mathbf{k},\omega)$ is expressed as [11, 171, 172]

$$\begin{aligned} \epsilon^*(\mathbf{k},\omega) &= 1 + \frac{\mathbf{k}\cdot\mu^*(\mathbf{k},\omega)\cdot\mathbf{k}}{i\omega\chi_{bb}(\mathbf{k},0)}, \\ \frac{1}{\epsilon^*(\mathbf{k},\omega)} &= 1 - \frac{\mathbf{k}\cdot\mu^{\text{ext}}(\mathbf{k},\omega)\cdot\mathbf{k}}{i\omega\chi_{bb}(\mathbf{k},0)} = 1 - \frac{\chi_{bb}(\mathbf{k},\omega)}{\chi_{bb}(\mathbf{k},0)}. \end{aligned} \tag{A.9}$$

It is informative to associate the internal response, $\mu^*(\mathbf{k},\omega)$, with the perturbation to the gradient of a *local* field, $v_b^*(\mathbf{k})$, [11, 173]

$$\left\langle \widehat{j}_b(\mathbf{k},t) \right\rangle = -\operatorname{Re}\mu^*(\mathbf{k},\omega) i\mathbf{k} v_b^*(\mathbf{k}) e^{i\omega t}. \tag{A.10}$$

The local field can be written in terms of the external potential and of the shielding factor [11],

$$\begin{aligned} v_b^*(\mathbf{k},t) \equiv \frac{1}{\epsilon^*(\mathbf{k},\omega)} v_b^{\text{ext}}(\mathbf{k},t) &= v_b^{\text{ext}}(\mathbf{k},t) - \frac{\chi_{bb}(\mathbf{k},\omega)}{\chi_{bb}(\mathbf{k},0)} v_b^{\text{ext}}(\mathbf{k},t) \\ &= v_b^{\text{ext}}(\mathbf{k},t) + \frac{\left\langle \widehat{b}(\mathbf{k},t) \right\rangle}{\chi_{bb}(\mathbf{k},0)}. \end{aligned} \tag{A.11}$$

In systems with short-range interactions $\chi^{-1}_{BB}(\mathbf{k},0)$ is not singular at $\mathbf{k}=0$ —it is the inverse of a thermodynamic derivative— and therefore, in the $\mathbf{k}\to 0$ and *then* $\omega\to 0$ limit, $\epsilon^*(\mathbf{k},\omega)\to 1$ [11, 13].

B

MARTIN'S ANALYSIS ON RESPONSES TO ELECTRIC PERTURBATIONS

Principles of dielectric response in charged systems

I review Martin's work on the interplay between responses to internal and external fields, presented in Ref. [86]. As discussed in Sec. 2.4, only external periodic density modulations are compatible with PBC. In spatially translational invariant systems, I write:

$$j_\alpha(\mathbf{k},\omega) = \sum_\beta \sigma_{\alpha\beta}(\mathbf{k},\omega) E_\beta(\mathbf{k},\omega), \tag{B.1}$$

where J_α and E_β are any components of the electric current and total field, respectively. Eq. (B.1) is decomposed in its transverse and longitudinal components,

$$\begin{aligned} \boldsymbol{j}^T(\mathbf{k},\omega) &= \sigma^T(\mathbf{k},\omega)\boldsymbol{E}^T(\mathbf{k},\omega) \\ \boldsymbol{j}^L(\mathbf{k},\omega) &= \sigma^L(\mathbf{k},\omega)\boldsymbol{E}^L(\mathbf{k},\omega). \end{aligned} \tag{B.2}$$

The ratio $\sigma_{\alpha\beta}(\mathbf{k},\omega) = j_\alpha(\mathbf{k},\omega)/E_\beta(\mathbf{k},\omega)$ is non analytic if a field $E_\beta(\mathbf{k},t)$ exists such that its Fourier transform $E_\beta(\mathbf{k},\omega_0) = 0$ for some ω_0. It can be argued that there always exists a field $E_\beta(\mathbf{k},t)$ which vanishes at $t < 0$ and whose Fourier transform is non zero for every ω_0 [86]. As discussed by Martin in Ref. [86], this reasoning does not apply to the total field $E_\beta(\mathbf{k},t)$, as it depends on the internal degrees of freedom and therefore, is not a tunable parameter [37, 86]. From these considerations it follows that, in response to a wave-vector dependent perturbation, the ratio is only measurable for a range of frequencies. $\sigma(\mathbf{k},\omega)$ and its analytic continuation do not satisfy the Kramers Kronig (KK) relations [86].

The Poisson and the continuity equations constrain severely the longitudinal component of the electric conductivity and Dielectric Function (DF) [86]:

$$\left[\omega + 4\pi i \sigma^L(\mathbf{k},\omega)\right] \boldsymbol{E}^L(\mathbf{k},\omega) = 0 \tag{B.3}$$

and the longitudinal DF can be written as

$$\begin{aligned} \epsilon^L(\mathbf{k},\omega) \equiv \frac{\mathbf{k}\cdot\boldsymbol{D}(\mathbf{k},\omega)}{\mathbf{k}\cdot\boldsymbol{E}(\mathbf{k},\omega)} &= \frac{\rho_{ext}(\mathbf{k},\omega)}{\rho_{tot}(\mathbf{k},\omega)} \\ &= 1 + 4\pi i \frac{\sigma^L(\mathbf{k},\omega)}{\omega}, \end{aligned} \tag{B.4}$$

ρ and ρ_{ext} being the total and external charge densities. Eq. (B.4) can be reformulated in terms of the charge response function, $\chi_{\rho\rho}$, [86]

$$\left[\epsilon^L(\mathbf{k},\omega)\right]^{-1} = 1 + 4\pi e^2 \frac{\chi_{\rho\rho}(\mathbf{k},\omega)}{k^2}. \tag{B.5}$$

Ref. [37] proves that the requirement for $\epsilon^L(\mathbf{k},\omega)$ to be analytic is:

$$\epsilon^L(\mathbf{k},0) = \left(1 + 4\pi e^2 \frac{\chi_{\rho\rho}(\mathbf{k})}{k^2}\right)^{-1} > 0 \tag{B.6}$$

Nevertheless, the only constrain for $\chi_{\rho\rho}$, Eq.(B.5), is [171]:

$$\begin{aligned} &\chi_{\rho\rho}(\mathbf{k}) < 0 \\ &\epsilon^L(\mathbf{k})^{-1} = 1 + 4\pi e^2 \frac{\chi_{\rho\rho}(\mathbf{k})}{k^2} \leqslant 1. \end{aligned} \tag{B.7}$$

In a seminal paper [171], Maximov demonstrates that, at finite wavelength, a negative dielectric constant is possible [86, 171–181].

The response function describes the reaction of the system to an external perturbation, it is causal and fulfills the KK relations [86, 171]. The *screened* response to the total charge density is

$$\rho(\mathbf{k},\omega) = \frac{4\pi e}{k^2}\chi^{sc}_{\rho\rho}(\mathbf{k},\omega)\left(\rho_{ext}(\mathbf{k},\omega) + \rho(\mathbf{k},\omega)\right). \tag{B.8}$$

Thanks to relation between the screened and unscreened response function [37, 86], one has

$$\chi_{\rho\rho}(\mathbf{k},\omega) = \frac{\chi^{sc}_{\rho\rho}(\mathbf{k},\omega)}{1 - \frac{4\pi e^2}{k^2}\chi^{sc}_{\rho\rho}(\mathbf{k},\omega)}. \tag{B.9}$$

The longitudinal DF, to be compared with Eq. (B.5), reads

$$\epsilon^L(\mathbf{k},\omega) = 1 - \frac{4\pi e^2}{k^2}\chi^{sc}_{\rho\rho}(\mathbf{k},\omega). \tag{B.10}$$

I have computed the wave-vector dependent dielectric constant of water from a 1 ns-long microcanonical trajectory with 4352 `TIP4P/2005` molecules. The procedure for the generation of the trajectory is described in Sec. 2.4. The charge response function and the dielectric constant are computed according to Eq.s (2.64), (2.65) and (B.5). Fig. B.1 displays the unscreened charge response function (times a minus sign) and the Bayesian-regression analysis of Eq. (2.64). The results are in good agreement with previous calculations [172, 179, 182–184].

The charge response function crosses -1 at $k^\star \approx 3\text{Å}^{-1}$, and it reflects in a stiff rise of $\epsilon(\mathbf{k})$. When $\chi_{\rho\rho}(\mathbf{k}) \approx -1$, as in a metal, the induced charge density perfectly balances the external fluctuation. In this condition, $\rho = -\rho_{ext} \Rightarrow \rho_{tot}(\mathbf{r}) = 0$, the system displays local perfect screening. Recently, the effect of long-range interactions has also been studied in the context of machine learning potentials [185–187]. The conclusions drawn in Ref. [185–187] are in accordance with those discussed in this section.

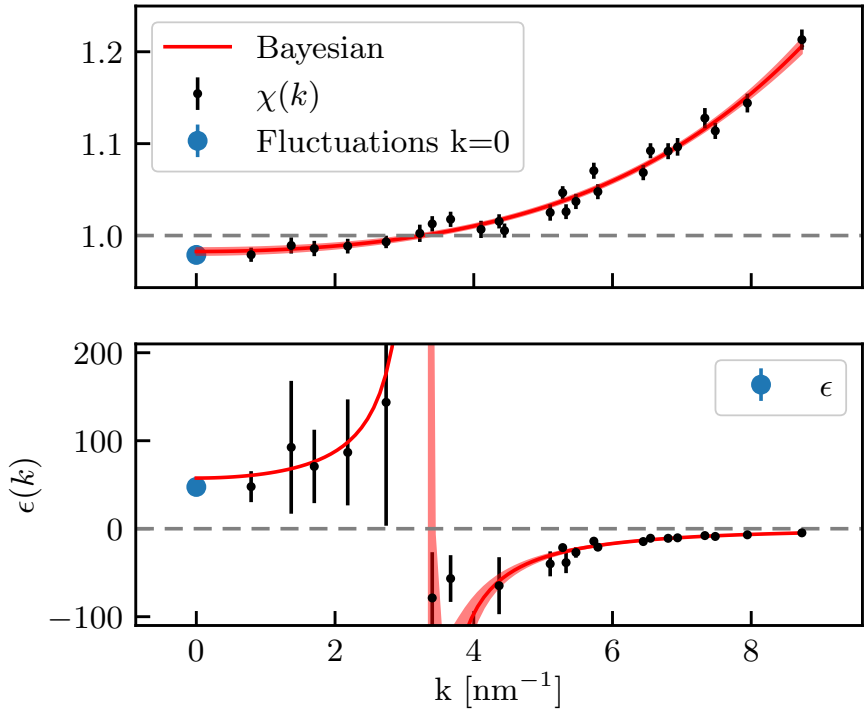

Figure B.1: Upper panel: Bayesian regression analysis of Eqs. (2.64)-(2.65). Lower panel: Wave-vector dependent dielectric constant, Eq. (B.5).The black dots are computed from a EMD trajectory of liquid `TIP4P/2005` water at 350 K and 0 bar. The red solid line is the estimate of Eq. (B.5) from the Bayesian-inference regression of Eq. (2.64).

Principles of transport phenomena in charged systems

The Onsager equations are [86, 87]

$$\boldsymbol{J} = \sigma \boldsymbol{E} - K_{12}/T \boldsymbol{\nabla} T$$
$$\boldsymbol{J}_q = K_{12} e \boldsymbol{E} - L_{qq} \boldsymbol{\nabla} T,$$

where $\boldsymbol{J}$ is the electric current, $\boldsymbol{J}_q$ is the heat current, $\boldsymbol{E}$ is the electric field, ∇T is the temperature gradient and, $\{\sigma, K_{12}, L_{qq}\}$ are the Onsager coefficients. In open circuit conditions, i. e. in a closed system, the charge density response to a temperature inhomogeneity is [12]

$$\delta\rho \equiv \chi_{\rho q} \delta T/T \tag{B.11}$$
$$\delta \boldsymbol{E} = \chi_{\rho q}/k^2 \boldsymbol{\nabla} T/T \tag{B.12}$$

where q is the heat density defined in Eq. (1.62) [21–24].

In order to determine the general relation between responses to external and internal fields, I observe that:

$$\delta n_\alpha(\mathbf{r}) = \int d\mathbf{r} \left(\chi_{AA}(\mathbf{r}, \mathbf{r}') \delta v_A(\mathbf{r}') + \chi_{AB}(\mathbf{r}, \mathbf{r}') \delta v_B(\mathbf{r}') + \chi_{A\rho}(\mathbf{r}, \mathbf{r}') \delta V_0(\mathbf{r}') \right) \tag{B.13}$$

where $v_{A,B}$ is the potential coupled to the respective density $n_{A,B}$ and V_0 is the Coulomb potential generated by the external charge density. I want to rewrite the response in terms of the total potential $V = V_0 + V_{int}$. V_{int} is the field of the induced charge $\langle \delta\rho \rangle$. V depends self-consistently on the induced charge as

$$\delta V_0 = \frac{\delta V_0}{\delta V}\left(\delta V + \frac{\delta V}{\delta v_B}\delta v_B + \frac{\delta V}{\delta v_A}\delta v_A\right) \tag{B.14}$$

$$= \frac{\delta V_0}{\delta V}\left(\delta V + \frac{\delta V_{int}}{\delta v_B}\delta v_B + \frac{\delta V_{int}}{\delta v_A}\delta v_A\right) \tag{B.15}$$

The response of the density, $n_\alpha(\mathbf{r})$, to the screened potential is

$$\begin{aligned}\delta n_\alpha(\mathbf{r}) = &\int d\mathbf{r}\,\chi_{AA}(\mathbf{r},\mathbf{r}')\delta v_A(\mathbf{r}') + \int d\mathbf{r}\,\chi_{AB}(\mathbf{r},\mathbf{r}')\delta v_B(\mathbf{r}') + \\ &\int d\mathbf{r}\,\chi_{A\rho}(\mathbf{r},\mathbf{r}')\left(\frac{\delta V_0}{\delta V}\left(\delta V(\mathbf{r}') + \frac{\delta V_{int}}{\delta v_B}\delta v_B(\mathbf{r}') + \frac{\delta V_{int}}{\delta v_A}\delta v_A(\mathbf{r}')\right)\right)\end{aligned} \tag{B.16}$$

Taking into account the total electrostatic potential V, all response functions change [86]:

$$\left.\frac{\delta n_A}{\delta v_B}\right|^{sc} \equiv \chi_{AB}{}^{sc} = \chi_{AB} + \chi_{A\rho}\frac{\delta V_0}{\delta V}\frac{\delta V_{int}}{\delta v_B} = \chi_{AB} + \chi_{A\rho}\frac{\epsilon^L}{k^2}\chi_{\rho B} \tag{B.17}$$

The thermo-power transport coefficient, K_{12}, is determined knowing that,

$$\sigma^L = \lim_{\omega\to 0}\lim_{\mathbf{k}\to 0}\omega\epsilon(\mathbf{k},\omega)/i \tag{B.18}$$

$$= \lim_{\omega\to 0}\lim_{\mathbf{k}\to 0}\chi_{\rho\rho}{}''^{sc}(\mathbf{k},\omega)\omega/k^2 \tag{B.19}$$

and therefore [86],

$$K_{12} = \lim_{\omega\to 0}\lim_{\mathbf{k}\to 0}\chi_{\rho q}(\mathbf{k},\omega)\omega\epsilon^L(\mathbf{k},\omega)/ik^2 \tag{B.20}$$

$$= \lim_{\omega\to 0}\lim_{\mathbf{k}\to 0}\chi_{\rho q}{}^{sc}\omega/ik^2 \tag{B.21}$$

In practical microcanonical simulations it is customary to set the total charge to zero, i. e. keeping $\mathbf{E} = 0$. This prescription is equivalent to the computation of the screened response function.

Hydrodynamic behavior of number and charge densities

Number and charge densities behave differently in the hydrodynamic regime [173, 188]. Their correlation function in the hydrodynamic regime can be approximated with a Lorentzian of width $\Gamma(\mathbf{k})$:

$$S(\mathbf{k},\omega) \approx \frac{S(\mathbf{k})\Gamma(\mathbf{k})}{\omega^2 + (\Gamma(\mathbf{k}))^2} \tag{B.22}$$

I compute the correlation functions from a 0.5 ns-long microcanonical simulation of molten CsF at 1200 K and 0 bar with 512 atoms. The details of the

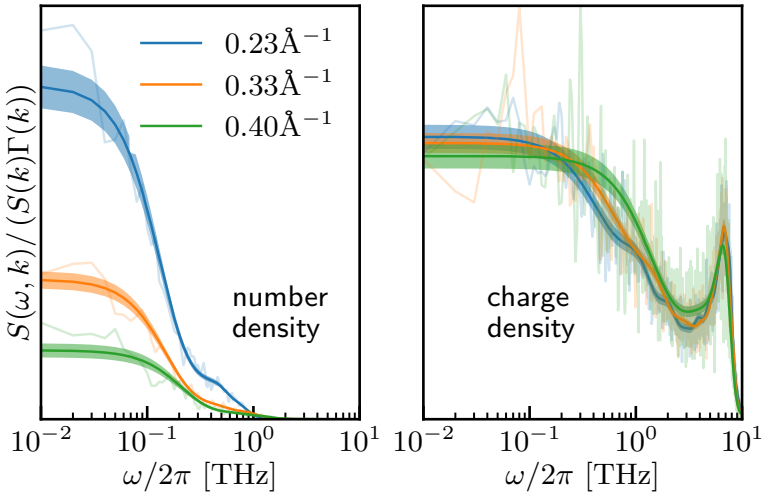

Figure B.2: Number (right panel) and charge (left panel) density correlation functions, Eq. (B.28), of molten CsF at 1200 K and 0 bar computed from a 0.5 ns-long microcanonical simulation. The shaded lines represent the raw data, the solid lines are the cepstral filtered signals.

simulation are described in Sec. 2.1. The number density, n, is a diffusive process. In the hydrodynamic regime, it follows the Fick's law. For the number correlation function, one has: [14, 173, 188]

$$S_{NN}(\mathbf{k}, \omega) \approx \frac{S_{NN}(\mathbf{k}) D k^2}{\omega^2 + (D k^2)^2}. \tag{B.23}$$

and $\Gamma_{NN} = D k^2$.

The dynamics of the charge density is strikingly different. Thanks to the continuity equation, I write [188]

$$-i\omega\rho(\mathbf{k}, \omega) = \rho(\mathbf{k}, t = 0) + i\mathbf{k} \cdot \boldsymbol{J}(\mathbf{k}, \omega) \tag{B.24}$$

$$-i\mathbf{k} \cdot \boldsymbol{E}(\mathbf{k}, \omega) = 4\pi\rho(\mathbf{k}, \omega) \tag{B.25}$$

$$\rho(\mathbf{k}, \omega) = \frac{\rho(\mathbf{k}, t = 0)}{-i\omega + 4\pi\sigma^{L}(\mathbf{k}, \omega)} \tag{B.26}$$

In the hydrodynamic regime, the charge correlation function is approximated as [173, 188]

$$S_{\rho\rho}(\mathbf{k}, \omega) \approx \frac{4\pi\sigma S_{\rho\rho}(\mathbf{k})}{\omega^2 + (4\pi\sigma)^2} \tag{B.27}$$

and $\Gamma_{\rho\rho} = 4\pi\sigma$. Charge fluctuations in the hydrodynamic regime are not propagating.

On one hand, as a function of frequency, $S_{NN}(\mathbf{k}, \omega)$ is a Lorentzian centered at 0 with width $D k^2$. On the other hand, the width of $S_{\rho\rho}(\mathbf{k}, \omega)$ is constant in the probed wave-vector [188]. In Fig. B.2 I show

$$\frac{S(\mathbf{k}, \omega)}{S(\mathbf{k})\Gamma(\mathbf{k})} \approx \frac{1}{\omega^2 + (\Gamma(\mathbf{k}))^2} \tag{B.28}$$

for number and charge density fluctuations. In the 0-frequency limit Eq. (B.28) is the inverse squared width of the Lorentzian. In the case of number fluctuations, the width shrinks a the wavelength increases: indeed, number density is a diffusive process. On the contrary, the width of the charge density correlation function is constant: in fact, charge fluctuations are neutralized in a microscopic time. Recently, the static effect of the long-range Coulomb interaction on charge density fluctuations has drawn some attention also in the machine learning community [185–187].

C

DEBENEDETTI CALCULATION OF PARTIAL MOLAR QUANTITIES

PARTIAL MOLAR enthalpies are thermodynamic derivatives, fundamental in the definition of heat density and current. Therefore, thermoelectric transport coefficients, such as the Seebeck and the Peltier coefficients, depend on these quantities [21, 46, 75, 87]. Indeed, convective invariance [15, 19], which holds for thermal conductivity, does not apply to off-diagonal Onsager coefficients [21]. In a series of seminal papers, Debenedetti set theoretical and numerical grounds for the computation of partial molar enthalpies from EMD simulation [22–24]. The estimation of these thermodynamic derivatives is rooted in the Kirkwood-Buff theory of fluctuations [189].

Let us define partial molar quantities, $\overline{\lambda}_i$, as[23]

$$\overline{\lambda}_i = \left.\frac{\partial \lambda}{\partial N_i}\right|_{T,P,[N]_i} \tag{C.1}$$

where $[N]_i$ indicates that I am keeping the number of all species fixed but the i-th. $\overline{\lambda}_i$ can be computed by finite differences evaluating

$$\lambda(N_1 \dots N_i \dots N_n, T, P) - \lambda(N_1 \dots N_i - 1 \dots N_n, T, P) \tag{C.2}$$

where n is the total number of species. Of course, finite differences approaches suffer from major numerical challenges and, because of electro-neutrality, are applicable only to neutral systems. Therefore, a general fluctuation procedure is extremely desirable. The partial molar energy can be written as:

$$\overline{E_j} = \frac{E}{V}\overline{V_j} + \sum_i \langle \delta E \delta N_i \rangle \left.\frac{\partial \beta \mu_i}{\partial N_i}\right|_{T,P,[N]_j} \tag{C.3}$$

$$\left.\frac{\partial E}{\partial \beta \mu_i}\right|_{\beta,[\beta\mu]_i} = \frac{E}{V} \tag{C.4}$$

$$\left.\frac{\partial E}{\partial V}\right|_{\beta,[\beta\mu]_i} = \langle \delta E \delta N_i \rangle \tag{C.5}$$

where $\langle \circ \rangle$ is the expectation value in the grand-canonical ensemble. Notably, Ref. [22, 189] prove that Eq (C.3) can be rewritten in terms of Kirkwood-Buff concentrations fluctuations:

$$\overline{E_j} = \frac{u \sum_i x_i |\Delta|_{ij}}{\sum_{ik} x_k x_i |\Delta|_{ik}} + \sum_i \langle \delta E \delta N_i \rangle \left(\frac{|\Delta|_{ij}}{|\Delta|} - \frac{\left(\sum_k x_k |\Delta|_{kj}\right)\left(\sum_k x_k |\Delta|_{ki}\right)}{\sum_{kl} x_k x_l |\Delta|_{kl}} |\Delta| \right) \tag{C.6}$$

where x is the mole fraction, $|\Delta|_{ij}$ it the cofactor of ij-th element in the determinant $|\Delta|$ of the matrix $\Delta_{ij} = \langle N_i N_j \rangle$.

The expectation value over the open control volume is estimated from a simple microcanonical trajectory, identifying 3 independent slices which divide the box in half in the xx, yy and zz directions. The aforementioned numerical protocol forces the molar quantities to satisfy the homogeneity constraint. The fluctuations method allows to compute thermodynamic derivatives, such as the partial molar quantities, from simple microcanonical simulations. Of course, fluctuations are evidently affected by ergodicity issues in solids and viscous low-temperature fluids. The method is particularly suitable for liquids well above the melting temperature.

D

VALIDATION OF THE MACHINE LEARNING INTERATOMIC POTENTIAL OF SOLID-STATE ELECTROLYTES

We expanded an existing dataset [55, 56] for LPS developed for Gaussian Approximation Potentials [190] The new structures are generated by means of the active learning scheme implemented in DP-GEN v0.12.0, [191]. We apply the scheme to reach around 10000 structures of α, β and γ-LPS, with temperatures ranging from 300 K to approximately 1000 K. The inter-atomic interaction is fitted via the version 4 NEP using GPUMD v3.9.4 [114], trained on Density Functional Theory (DFT) energies and forces, estimated at the PBEsol level of theory [116, 192]. The QUANTUM ESPRESSO™ code was used for all DFT single point calculations [193–196]. The electronic Kohn-Sham wave-function and density are computed with 70 Ry and 400 Ry cutoffs respectively, PAW pseudo-potentials and 0.008 Ry Marzari-Vanderbilt smearing [197, 198]. The NEP is tested on a separate test-set containing 2600 configurations [45]. Fig. D.1 displays the NEP performance on the test-set.

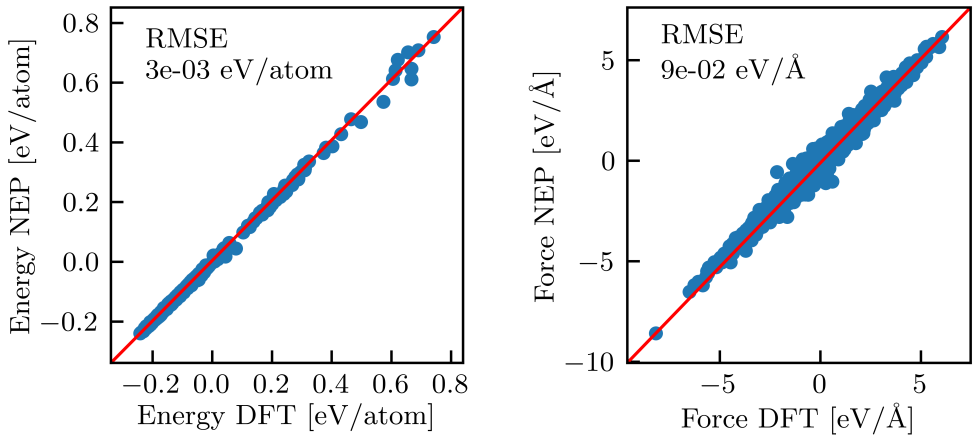

Figure D.1: Energy (left panel) and force (right panel) parity plots of NEP predictions vs DFT reference data.

ACKNOWLEDGMENTS

What an adventure! I shared my PhD with amazing people. To you, I want to express my deepest gratitude.

To my supervisor, Stefano Baroni, for his invaluable mentorship and for the patience he showed with me. His passion for science and his work ethic have inspired my PhD journey.

To all my friends, you filled my years in Trieste with unforgettable and precious memories. To Alfredo, for your friendship and our countless chats, even the ones about physics; to Giacomo, Cesare, Martina, Pietro, Michele, and Anna, for a decade-long friendship during the bachelor (and even before!), master and PhD years; to Paolo, Davide, and Federico, for our conversations and our absurd, irresistible laughter.

To my family, especially my mum and dad, for your constant and unwavering support, you are the core of my development as a person. There is too much to say, yet I struggle to find the right words. Thank you

BIBLIOGRAPHY

[1] Herbert B Callen. *Thermodynamics and an introduction to thermostatistics; 2nd ed.* New York, NY: Wiley, 1985. URL: https://cds.cern.ch/record/450289.

[2] L. Onsager and S. Machlup. "Fluctuations and Irreversible Processes." In: *Phys. Rev.* 91 (6 1953), pp. 1505–1512. DOI: 10.1103/PhysRev.91.1505.

[3] H. B. G. Casimir. "On Onsager's Principle of Microscopic Reversibility." In: *Rev. Mod. Phys.* 17 (2-3 1945), pp. 343–350. DOI: 10.1103/RevModPhys.17.343.

[4] Lars Onsager. "Reciprocal Relations in Irreversible Processes. I." In: *Phys. Rev.* 37 (4 1931), pp. 405–426. DOI: 10.1103/PhysRev.37.405.

[5] Lars Onsager. "Reciprocal Relations in Irreversible Processes. II." In: *Phys. Rev.* 38 (12 1931), pp. 2265–2279. DOI: 10.1103/PhysRev.38.2265.

[6] Ryogo Kubo. "Statistical-Mechanical Theory of Irreversible Processes. I. General Theory and Simple Applications to Magnetic and Conduction Problems." In: *J. Phys. Soc. Jpn.* 12.6 (1957), pp. 570–586. DOI: 10.1143/JPSJ.12.570.

[7] Ryogo Kubo, Mario Yokota, and Sadao Nakajima. "Statistical-Mechanical Theory of Irreversible Processes. II. Response to Thermal Disturbance." In: *J. Phys. Soc. Jpn.* 12.11 (1957), pp. 1203–1211. DOI: 10.1143/JPSJ.12.1203.

[8] Melville S. Green. "Markoff Random Processes and the Statistical Mechanics of Time-Dependent Phenomena." In: *The Journal of Chemical Physics* 20.8 (1952), pp. 1281–1295. ISSN: 0021-9606. DOI: 10.1063/1.1700722.

[9] Melville S. Green. "Markoff Random Processes and the Statistical Mechanics of Time-Dependent Phenomena. II. Irreversible Processes in Fluids." In: *The Journal of Chemical Physics* 22.3 (1954), pp. 398–413. ISSN: 0021-9606. DOI: 10.1063/1.1740082.

[10] John G. Kirkwood. "The Statistical Mechanical Theory of Transport Processes I. General Theory." In: *The Journal of Chemical Physics* 14.3 (1946), pp. 180–201. ISSN: 0021-9606. DOI: 10.1063/1.1724117.

[11] R Kubo. "The fluctuation-dissipation theorem." In: *Reports on Progress in Physics* 29.1 (1966), p. 255. DOI: 10.1088/0034-4885/29/1/306.

[12] J. M. Luttinger. "Theory of Thermal Transport Coefficients." In: *Phys. Rev.* 135 (6A 1964), A1505–A1514. DOI: 10.1103/PhysRev.135.A1505.

[13] Martin Paul C. *Statistical Mechanics of Equilibrium and Non-equilibrium*. Ed. by J. Meixner. North-Holland Publishing Company, 1965, pp. 100–125.

[14] D. Forster. *Hydrodynamic Fluctuations, Broken Symmetry, and Correlation Functions*. Frontiers in Physics : a lecture note and reprint series. W. A. Benjamin, Advanced Book Program, 1975. ISBN: 9780805325607.

[15] Stefano Baroni, Riccardo Bertossa, Loris Ercole, Federico Grasselli, and Aris Marcolongo. *Heat Transport in Insulators from Ab Initio Green-Kubo Theory*. Ed. by Wanda Andreoni and Sidney Yip. Cham: Springer International Publishing, 2018, pp. 1–36. ISBN: 978-3-319-50257-1. DOI: 10.1007/978-3-319-50257-1_12-1.

[16] Stefano Baroni. *Green and Kubo forge the arrow of time*. 2024. arXiv: 2404.09291 [cond-mat.stat-mech]. URL: https://arxiv.org/abs/2404.09291.

[17] A. Einstein. "Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen." In: *Annalen der Physik* 322.8 (1905), pp. 549–560. DOI: https://doi.org/10.1002/andp.19053220806.

[18] Eugene Helfand. "Transport Coefficients from Dissipation in a Canonical Ensemble." In: *Phys. Rev.* 119 (1 1960), pp. 1–9. DOI: 10.1103/PhysRev.119.1.

[19] Federico Grasselli and Stefano Baroni. "Invariance principles in the theory and computation of transport coefficients." In: *The European Physical Journal B* 94.8 (2021), p. 160. ISSN: 1434-6036. DOI: 10.1140/epjb/s10051-021-00152-5.

[20] Leo P Kadanoff and Paul C Martin. "Hydrodynamic equations and correlation functions." In: *Annals of Physics* 24 (1963), pp. 419–469. ISSN: 0003-4916. DOI: https://doi.org/10.1016/0003-4916(63)90078-2.

[21] S.R. de Groot and P. Mazur. *Non-equilibrium Thermodynamics*. Dover Books on Physics. Dover Publications, 1984. ISBN: 9780486647418.

[22] Pablo G. Debenedetti. "The statistical mechanical theory of concentration fluctuations in mixtures." In: *J. Chem. Phys.* 87.2 (1987), pp. 1256–1260. DOI: 10.1063/1.453308.

[23] Pablo G. Debenedetti. "Fluctuation-based computer calculation of partial molar properties. I. Molecular dynamics simulation of constant volume fluctuations." In: *J. Chem. Phys.* 86.12 (1987), pp. 7126–7137. DOI: 10.1063/1.452362.

[24] Pablo G. Debenedetti. "Fluctuation-based computer calculation of partial molar properties. II. A numerically accurate method for the determination of partial molar energies and enthalpies." In: *J. Chem. Phys.* 88.4 (1988), pp. 2681–2684. DOI: 10.1063/1.453996.

[25] Hazime Mori. "Statistical-Mechanical Theory of Transport in Fluids." In: *Phys. Rev.* 112 (6 1958), pp. 1829–1842. DOI: 10.1103/PhysRev.112.1829.

[26] Melville S. Green. "Comment on a Paper of Mori on Time-Correlation Expressions for Transport Properties." In: *Phys. Rev.* 119 (3 1960), pp. 829–830. DOI: 10.1103/PhysRev.119.829.

[27] Stephen Stackhouse, Lars Stixrude, and Bijaya B. Karki. "Thermal Conductivity of Periclase (MgO) from First Principles." In: *Phys. Rev. Lett.* 104 (20 2010), p. 208501. DOI: 10.1103/PhysRevLett.104.208501.

[28] Davide Tisi, Linfeng Zhang, Riccardo Bertossa, Han Wang, Roberto Car, and Stefano Baroni. "Heat transport in liquid water from first-principles and deep neural network simulations." In: *Phys. Rev. B* 104 (22 2021), p. 224202. DOI: 10.1103/PhysRevB.104.224202.

[29] Loris Ercole, Aris Marcolongo, Paolo Umari, and Stefano Baroni. "Gauge Invariance of Thermal Transport Coefficients." In: *Journal of Low Temperature Physics* 185.1 (2016), pp. 79–86. ISSN: 1573-7357. DOI: 10.1007/s10909-016-1617-6.

[30] Federico Grasselli and Stefano Baroni. "Topological quantization and gauge invariance of charge transport in liquid insulators." In: *Nature Physics* 15.9 (2019), pp. 967–972. ISSN: 1745-2481. DOI: 10.1038/s41567-019-0562-0.

[31] Riccardo Bertossa, Federico Grasselli, Loris Ercole, and Stefano Baroni. "Theory and Numerical Simulation of Heat Transport in Multicomponent Systems." In: *Phys. Rev. Lett.* 122 (25 2019), p. 255901. DOI: 10.1103/PhysRevLett.122.255901.

[32] Paolo Pegolo, Stefano Baroni, and Federico Grasselli. "Temperature- and vacancy-concentration-dependence of heat transport in Li3ClO from multi-method numerical simulations." In: *npj Computational Materials* 8.1 (2022), p. 24. ISSN: 2057-3960. DOI: 10.1038/s41524-021-00693-4.

[33] Federico Grasselli. "Investigating finite-size effects in molecular dynamics simulations of ion diffusion, heat transport, and thermal motion in superionic materials." In: *The Journal of Chemical Physics* 156.13 (Apr. 2022), p. 134705. ISSN: 0021-9606. DOI: 10.1063/5.0087382.

[34] Paolo Pegolo, Stefano Baroni, and Federico Grasselli. "Self-interaction and transport of solvated electrons in molten salts." In: *The Journal of Chemical Physics* 159.9 (Sept. 2023), p. 094116. ISSN: 0021-9606. DOI: 10.1063/5.0169474.

[35] Lars Stixrude, Stefano Baroni, and Federico Grasselli. "Thermal and Tidal Evolution of Uranus with a Growing Frozen Core." In: *The Planetary Science Journal* 2.6 (2021), p. 222. DOI: 10.3847/PSJ/ac2a47.

[36] Aris Marcolongo, Loris Ercole, and Stefano Baroni. "Gauge Fixing for Heat-Transport Simulations." In: *Journal of Chemical Theory and Computation* 16.5 (2020). PMID: 32239930, pp. 3352–3362. DOI: 10.1021/acs.jctc.9b01174.

[37] P. Nozieres and D. Pines. *Theory Of Quantum Liquids*. Advanced Books Classics. Avalon Publishing, 1999. ISBN: 9780813346533.

[38] N. Galamba, C. A. Nieto de Castro, and James F. Ely. "Equilibrium and nonequilibrium molecular dynamics simulations of the thermal conductivity of molten alkali halides." In: *J. Chem. Phys.* 126.20 (2007), p. 204511. ISSN: 0021-9606. DOI: 10.1063/1.2734965.

[39] Sara Bonella, Mauro Ferrario, and Giovanni Ciccotti. "Thermal Diffusion in Binary Mixtures: Transient Behavior and Transport Coefficients from Equilibrium and Nonequilibrium Molecular Dynamics." In: *Langmuir* 33.42 (2017). PMID: 28915729, pp. 11281–11290. DOI: 10.1021/acs.langmuir.7b02565.

[40] Fernando Bresme, Anders Lervik, Dick Bedeaux, and Signe Kjelstrup. "Water Polarization under Thermal Gradients." In: *Phys. Rev. Lett.* 101 (2 2008), p. 020602. DOI: 10.1103/PhysRevLett.101.020602.

[41] Silvia Di Lecce and Fernando Bresme. "Thermal Polarization of Water Influences the Thermoelectric Response of Aqueous Solutions." In: *The Journal of Physical Chemistry B* 122.5 (2018). PMID: 29293343, pp. 1662–1668. DOI: 10.1021/acs.jpcb.7b10960.

[42] Silvia Di Lecce, Tim Albrecht, and Fernando Bresme. "A computational approach to calculate the heat of transport of aqueous solutions." In: *Scientific Reports* 7.1 (2017), p. 44833. ISSN: 2045-2322. DOI: 10.1038/srep44833.

[43] Marcello Puligheddu and Giulia Galli. "Atomistic simulations of the thermal conductivity of liquids." In: *Phys. Rev. Mater.* 4 (5 2020), p. 053801. DOI: 10.1103/PhysRevMaterials.4.053801.

[44] Loris Ercole, Aris Marcolongo, and Stefano Baroni. "Accurate thermal conductivities from optimally short molecular dynamics simulations." In: *Sci. Rep.* 7.1 (2017), p. 15835. ISSN: 2045-2322. DOI: 10.1038/s41598-017-15843-2.

[45] Paolo Pegolo, Enrico Drigo, Federico Grasselli, and Stefano Baroni. "Transport coefficients from equilibrium molecular dynamics". Submitted to: *The Journal of Chemical Physics*. 2024.

[46] Enrico Drigo, Stefano Baroni, and Paolo Pegolo. "Seebeck Coefficient of Ionic Conductors from Bayesian Regression Analysis." In: *Journal of Chemical Theory and Computation* 20.14 (2024). PMID: 38856670, pp. 6152–6159. DOI: 10.1021/acs.jctc.4c00124.

[47] "Ionic sizes and born repulsive parameters in the NaCl-type alkali halides—I: The Huggins-Mayer and Pauling forms." In: *J. Phys. Chem. Solids* 25.1 (1964), pp. 31–43. ISSN: 0022-3697. DOI: https://doi.org/10.1016/0022-3697(64)90159-3.

[48] "Ionic sizes and born repulsive parameters in the NaCl-type alkali halides—II: The generalized Huggins-Mayer form." In: *J. Phys. Chem. Solids* 25.1 (1964), pp. 45–52. ISSN: 0022-3697. DOI: https://doi.org/10.1016/0022-3697(64)90160-X.

[49] Jia Wang, Ze Sun, Guimin Lu, and Jianguo Yu. "Molecular Dynamics Simulations of the Local Structures and Transport Coefficients of Molten Alkali Chlorides." In: *J. Phys. Chem. B* 118.34 (2014). PMID: 25105467, pp. 10196–10206. DOI: 10.1021/jp5050332.

[50] A. P. Thompson et al. "LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales." In: *Comp. Phys. Comm.* 271 (2022), p. 108171. DOI: 10.1016/j.cpc.2021.108171.

[51] Norbert Wiener. "Generalized harmonic analysis." In: *Acta Mathematica* 55.none (1930), pp. 117 –258. DOI: 10.1007/BF02546511.

[52] A. Khintchine. "Korrelationstheorie der stationären stochastischen Prozesse." In: *Mathematische Annalen* 109.1 (1934), pp. 604–615. ISSN: 1432-1807. DOI: 10.1007/BF01449156.

[53] Eric W. Weisstein. https://mathworld.wolfram.com/PolygammaFunction.html. *Polygamma Function*. From MathWorld–A Wolfram Web Resource.

[54] Federico Grasselli, Lars Stixrude, and Stefano Baroni. "Heat and charge transport in H2O at ice-giant conditions from ab initio molecular dynamics simulations." In: *Nature Communications* 11.1 (2020), p. 3605. ISSN: 2041-1723. DOI: 10.1038/s41467-020-17275-5.

[55] Davide Tisi, Federico Grasselli, Lorenzo Gigli, and Michele Ceriotti. "Thermal conductivity of Li_3PS_4 solid electrolytes with ab initio accuracy." In: *Phys. Rev. Mater.* 8 (6 2024), p. 065403. DOI: 10.1103/PhysRevMaterials.8.065403.

[56] Lorenzo Gigli, Davide Tisi, Federico Grasselli, and Michele Ceriotti. "Mechanism of Charge Transport in Lithium Thiophosphate." In: *Chemistry of Materials* 36.3 (2024), pp. 1482–1496. DOI: 10.1021/acs.chemmater.3c02726.

[57] Cesare Malosso, Linfeng Zhang, Roberto Car, Stefano Baroni, and Davide Tisi. "Viscosity in water from first-principles and deep-neural-network simulations." In: *Npj Comput. Mater.* 8.1 (2022). ISSN: 2057-3960. DOI: 10.1038/s41524-022-00830-7.

[58] S. S. Shapiro and M. B. Wilk. "An analysis of variance test for normality (complete samples)." In: *Biometrika* 52.3-4 (1965), pp. 591–611. ISSN: 0006-3444. DOI: 10.1093/biomet/52.3-4.591.

[59] Paolo Pegolo. "Charge and heat transport in ionic conductors." PhD thesis. SISSA, 2023. DOI: 20.500.11767/130672.

[60] Scott Nestler and Andrew Hall. "The Variance Gamma Distribution." In: *Significance* 16.5 (2019), pp. 10–11. ISSN: 1740-9705. DOI: 10.1111/j.1740-9713.2019.01314.x.

[61] Christopher M. Bishop. *Pattern Recognition and Machine Learning*. 1st ed. Springer, 2006. Chap. 3.

[62] A. W. van der Vaart. *Asymptotic Statistics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1998.

[63] C. De Boor. *A Practical Guide to Splines*. Applied Mathematical Sciences. Springer New York, 1978. ISBN: 9783540903567.

[64] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman. "emcee: The MCMC Hammer." In: *PASP* 125 (2013), pp. 306–312. DOI: 10.1086/670067.

[65] Jonathan Goodman and Jonathan Weare. "Ensemble samplers with affine invariance." In: *Comm. App. Math. Comp. Sci.* 5.1 (2010), pp. 65–80.

[66] J. L. F. Abascal and C. Vega. "A general purpose model for the condensed phases of water: TIP4P/2005." In: *The Journal of Chemical Physics* 123.23 (2005), p. 234505. ISSN: 0021-9606. DOI: 10.1063/1.2121687.

[67] Jeremy C. Palmer, Rakesh S. Singh, Renjie Chen, Fausto Martelli, and Pablo G. Debenedetti. "Density and bond-orientational relaxations in supercooled water." In: *Molecular Physics* 114.18 (2016), pp. 2580–2585. DOI: 10.1080/00268976.2016.1179351.

[68] Juan L. Aragones Helena L. Pi, Carlos Vega, Jose L.F. Abascal Eva G. Noya, Miguel A. Gonzalez, and Carl McBride. "Anomalies in water as obtained from computer simulations of the TIP4P/2005 model: density maxima, and density, isothermal compressibility and heat capacity minima." In: *Molecular Physics* 107.4-6 (2009), pp. 365–374. DOI: 10.1080/00268970902784926.

[69] Handle Philip H. and Sciortino Francesco. "Potential energy landscape of TIP4P/2005 water." In: *The Journal of Chemical Physics* 148.13 (2018), p. 134505. ISSN: 0021-9606. DOI: 10.1063/1.5023894.

[70] Philip H. Handle and Francesco Sciortino. "The Adam–Gibbs relation and the TIP4P/2005 model of water." In: *Molecular Physics* 116.21-22 (2018). PMID: 30338319, pp. 3366–3371. DOI: 10.1080/00268976.2018.1471230.

[71] Sciortino Francesco, Gartner Thomas E. III, and Pablo G. Debenedetti. "Free-energy landscape and spinodals for the liquid–liquid transition of the TIP4P/2005 and TIP4P/Ice models of water." In: *The Journal of Chemical Physics* 160.10 (2024), p. 104501. ISSN: 0021-9606. DOI: 10.1063/5.0196964.

[72] Pablo G. Debenedetti, Francesco Sciortino, and Gül H. Zerze. "Second critical point in two realistic models of water." In: *Science* 369.6501 (2020), pp. 289–292. DOI: 10.1126/science.abb9796.

[73] Amir Haji-Akbari and Pablo G. Debenedetti. "Direct calculation of ice homogeneous nucleation rate for a molecular model of water." In: *Proceedings of the National Academy of Sciences* 112.34 (2015), pp. 10582–10588. DOI: 10.1073/pnas.1509267112.

[74] M.P. Allen and D.J. Tildesley. *Computer Simulation of Liquids*. OUP Oxford, 2017. Chap. 2. ISBN: 9780192524706. URL: https://books.google.it/books?id=WFExDwAAQBAJ.

[75] Enrico Drigo and Stefano Baroni. "Seebeck Coefficient of Liquid Water from Equilibrium Molecular Dynamics." In: *Journal of Chemical Theory and Computation* 19.23 (2023). PMID: 37823575, pp. 8855–8860. DOI: 10.1021/acs.jctc.3c00760.

[76] P. Wirnsberger, D. Fijan, A. Šarić, M. Neumann, C. Dellago, and D. Frenkel. "Non-equilibrium simulations of thermally induced electric fields in water." In: *The Journal of Chemical Physics* 144.22 (2016), p. 224102. ISSN: 0021-9606. DOI: 10.1063/1.4953036.

[77] Peter Wirnsberger, Domagoj Fijan, Roger A. Lightwood, Anđela Šarić, Christoph Dellago, and Daan Frenkel. "Numerical evidence for thermally induced monopoles." In: *Proceedings of the National Academy of Sciences* 114.19 (2017), pp. 4911–4914. DOI: 10.1073/pnas.1621494114. URL: https://www.pnas.org/doi/abs/10.1073/pnas.1621494114.

[78] P. Wirnsberger, C. Dellago, D. Frenkel, and A. Reinhardt. "Theoretical Prediction of Thermal Polarization." In: *Phys. Rev. Lett.* 120 (22 2018), p. 226001. DOI: 10.1103/PhysRevLett.120.226001.

[79] Jeff Armstrong, Anders Lervik, and Fernando Bresme. "Enhancement of the Thermal Polarization of Water via Heat Flux and Dipole Moment Dynamic Correlations." In: *The Journal of Physical Chemistry B* 117.47 (2013). PMID: 24168463, pp. 14817–14826. DOI: 10.1021/jp408485d.

[80] Jeff Armstrong and Fernando Bresme. "Temperature inversion of the thermal polarization of water." In: *Phys. Rev. E* 92 (6 2015), p. 060103. DOI: 10.1103/PhysRevE.92.060103.

[81] Aidan Chapman and Fernando Bresme. "Polarisation of water under thermal fields: the effect of the molecular dipole and quadrupole moments." In: *Phys. Chem. Chem. Phys.* 24 (24 2022), pp. 14924–14936. DOI: 10.1039/D2CP00756H.

[82] Frank Römer, Fernando Bresme, Jordan Muscatello, Dick Bedeaux, and J. Miguel Rubí. "Thermomolecular Orientation of Nonpolar Fluids." In: *Phys. Rev. Lett.* 108 (10 2012), p. 105901. DOI: 10.1103/PhysRevLett.108.105901.

[83] W. H. Furry, R. Clark Jones, and L. Onsager. "On the Theory of Isotope Separation by Thermal Diffusion." In: *Phys. Rev.* 55 (11 1939), pp. 1083–1095. DOI: 10.1103/PhysRev.55.1083.

[84] G. W. Stewart, D. O. Holland, and L. M. Reynolds. "Orientation of Liquid Crystals by Heat Conduction." In: *Phys. Rev.* 58 (2 1940), pp. 174–176. DOI: 10.1103/PhysRev.58.174.

[85] G. Jeffrey Snyder and Eric S. Toberer. "Complex thermoelectric materials." In: *Nature Materials* 7.2 (2008), pp. 105–114. ISSN: 1476-4660. DOI: 10.1038/nmat2090.

[86] Paul C. Martin. "Sum Rules, Kramers-Kronig Relations, and Transport Coefficients in Charged Systems." In: *Phys. Rev.* 161 (1 1967), pp. 143–155. DOI: 10.1103/PhysRev.161.143.

[87] Herbert B. Callen. "The Application of Onsager's Reciprocal Relations to Thermoelectric, Thermomagnetic, and Galvanomagnetic Effects." In: *Phys. Rev.* 73 (11 1948), pp. 1349–1358. DOI: 10.1103/PhysRev.73.1349.

[88] Aris Marcolongo, Paolo Umari, and Stefano Baroni. "Microscopic theory and quantum simulation of atomic heat transport." In: *Nature Physics* 12.1 (2016), pp. 80–84. ISSN: 1745-2481. DOI: 10.1038/nphys3509.

[89] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma. "The missing term in effective pair potentials." In: *J. Phys. Chem.* 91.24 (1987), pp. 6269–6271. DOI: 10.1021/j100308a038.

[90] Dominique Chrétien, Paule Bénit, Hyung-Ho Ha, Susanne Keipert, Riyad El-Khoury, Young-Tae Chang, Martin Jastroch, Howard T. Jacobs, Pierre Rustin, and Malgorzata Rak. "Mitochondria are physiologically maintained at close to 50 °C." In: *PLOS Biology* 16.1 (2018), pp. 1–17. DOI: 10.1371/journal.pbio.2003992.

[91] David Macherel, Francis Haraux, Hervé Guillou, and Olivier Bourgeois. "The conundrum of hot mitochondria." In: *Biochimica biophysica acta (BBA) - Bioenergetics* 1862.2 (2021), p. 148348. DOI: 10.1016/j.bbabio.2020.148348.

[92] Xiangjun Di, Dejiang Wang, Qian Peter Su, Yongtao Liu, Jiayan Liao, Mahnaz Maddahfar, Jiajia Zhou, and Dayong Jin. "Spatiotemporally mapping temperature dynamics of lysosomes and mitochondria using cascade organelle-targeting upconversion nanoparticles." In: *Proceedings of the National Academy of Sciences* 119.45 (2022), e2207402119. DOI: 10.1073/pnas.2207402119.

[93] Alois Würger. "Thermopower of ionic conductors and ionic capacitors." In: *Phys. Rev. Res.* 2 (4 2020), p. 042030. DOI: 10.1103/PhysRevResearch.2.042030.

[94] André Luiz Sehnem and Mathijs Janssen. "On the time-dependent electrolyte Seebeck effect." In: *The Journal of Chemical Physics* 154.16 (2021), p. 164511. ISSN: 0021-9606. DOI: 10.1063/5.0045137.

[95] Michael P. Allen and Dominic J. Tildesley. "Nonequilibrium molecular dynamics." In: *Computer Simulation of Liquids*. Oxford University Press, 2017. ISBN: 9780198803195. DOI: 10.1093/oso/9780198803195.003.0011.

[96] Patrick K. Schelling, Simon R. Phillpot, and Pawel Keblinski. "Comparison of atomic-level simulation methods for computing thermal conductivity." In: *Phys. Rev. B* 65 (14 2002), p. 144306. DOI: 10.1103/PhysRevB.65.144306.

[97] Haikuan Dong, Zheyong Fan, Libin Shi, Ari Harju, and Tapio Ala-Nissila. "Equivalence of the equilibrium and the nonequilibrium molecular dynamics methods for thermal conductivity calculations: From bulk to nanowire silicon." In: *Phys. Rev. B* 97 (9 2018), p. 094305. DOI: 10.1103/PhysRevB.97.094305.

[98] Eric W. Weisstein. https://mathworld.wolfram.com/MovingAverage.html. *Moving Average*. From MathWorld–A Wolfram Web Resource.

[99] Han Wang, Linfeng Zhang, Jiequn Han, and Weinan E. "DeePMD-kit: A deep learning package for many-body potential energy representation and molecular dynamics." In: *Comp. Phys. Comm.* 228 (2018), pp. 178–184. ISSN: 0010-4655. DOI: https://doi.org/10.1016/j.cpc.2018.03.016.

[100] Linfeng Zhang, Jiequn Han, Han Wang, Roberto Car, and Weinan E. "Deep Potential Molecular Dynamics: A Scalable Model with the Accuracy of Quantum Mechanics." In: *Phys. Rev. Lett.* 120 (14 2018), p. 143001. DOI: 10.1103/PhysRevLett.120.143001.

[101] Wenshuo Liang, Guimin Lu, and Jianguo Yu. "Theoretical prediction on the local structure and transport properties of molten alkali chlorides by deep potentials." In: *J. Mater. Sci. Technol.* 75 (2021), pp. 78–85. ISSN: 1005-0302. DOI: https://doi.org/10.1016/j.jmst.2020.09.040.

[102] Giovanni Bussi, Davide Donadio, and Michele Parrinello. "Canonical sampling through velocity rescaling." In: *The Journal of Chemical Physics* 126.1 (2007), p. 014101. ISSN: 0021-9606. DOI: 10.1063/1.2408420.

[103] Paolo Pegolo, Stefano Baroni, and Federico Grasselli. "Topology, oxidation states, and charge transport in ionic conductors." In: *Ann. Phys.* (2022), p. 2200123.

[104] Paolo Pegolo, Federico Grasselli, and Stefano Baroni. "Oxidation States, Thouless' Pumps, and Nontrivial Ionic Transport in Nonstoichiometric Electrolytes." In: *Phys. Rev. X* 10 (4 2020), p. 041031. DOI: 10.1103/PhysRevX.10.041031.

[105] R. H. Detig and D. H. Archer. "Thermoelectric Effects in Fused Ionic Materials." In: *The Journal of Chemical Physics* 38.3 (1963), pp. 661–666. ISSN: 0021-9606. DOI: 10.1063/1.1733719.

[106] W. H. Furry, R. Clark Jones, and L. Onsager. "On the Theory of Isotope Separation by Thermal Diffusion." In: *Phys. Rev.* 55 (11 1939), pp. 1083–1095. DOI: 10.1103/PhysRev.55.1083.

[107] Filippo Bigi, Sanggyu Chong, Michele Ceriotti, and Federico Grasselli. "A prediction rigidity formalism for low-cost uncertainties in trained neural networks." In: *Machine Learning: Science and Technology* (2024). URL: http://iopscience.iop.org/article/10.1088/2632-2153/ad805f.

[108] Cong Li, Zhen yu Wang, Zhen jiang He, Yun jiao Li, Jing Mao, Ke hua Dai, Cheng Yan, and Jun chao Zheng. "An advance review of solid-state battery: Challenges, progress and prospects." In: *Sustainable Materials and Technologies* 29 (2021), e00297. ISSN: 2214-9937. DOI: https://doi.org/10.1016/j.susmat.2021.e00297.

[109] Ömer Ulaş Kudu et al. "Structural details in Li3PS4: Variety in thiophosphate building blocks and correlation to ion transport." In: *Energy Storage Materials* 44 (2022), pp. 168–179. ISSN: 2405-8297. DOI: https://doi.org/10.1016/j.ensm.2021.10.021.

[110] Kenji Homma, Masao Yonemura, Takeshi Kobayashi, Miki Nagao, Masaaki Hirayama, and Ryoji Kanno. "Crystal structure and phase transitions of the lithium ionic conductor Li3PS4." In: *Solid State Ionics* 182.1 (2011), pp. 53–58. ISSN: 0167-2738. DOI: https://doi.org/10.1016/j.ssi.2010.10.001.

[111] Takuya Kimura, Takeaki Inaoka, Ryo Izawa, Takumi Nakano, Chie Hotehama, Atsushi Sakuda, Masahiro Tatsumisago, and Akitoshi Hayashi. "Stabilizing High-Temperature alpha-Li3PS4 by Rapidly Heating the Glass." In: *Journal of the American Chemical Society* 145.26 (2023). PMID: 37340711, pp. 14466–14474. DOI: 10.1021/jacs.3c03827.

[112] Niek J.J. de Klerk, Eveline van der Maas, and Marnix Wagemaker. "Analysis of Diffusion in Solid-State Electrolytes through MD Simulations, Improvement of the Li-Ion Conductivity in beta-Li3PS4 as an Example." In: *ACS Applied Energy Materials* 1.7 (2018). PMID: 30057999, pp. 3230–3242. DOI: 10.1021/acsaem.8b00457.

[113] Frazer N. Forrester, James A. Quirk, Theodosios Famprikis, and James A. Dawson. "Disentangling Cation and Anion Dynamics in Li3PS4 Solid Electrolytes." In: *Chemistry of Materials* 34.23 (2022), pp. 10561–10571. DOI: 10.1021/acs.chemmater.2c02637.

[114] Zheyong Fan et al. "GPUMD: A package for constructing accurate machine-learned potentials and performing highly efficient atomistic simulations." In: *The Journal of Chemical Physics* 157.11 (2022), p. 114801. ISSN: 0021-9606. DOI: 10.1063/5.0106617.

[115] Zheyong Fan, Luiz Felipe C. Pereira, Hui-Qiong Wang, Jin-Cheng Zheng, Davide Donadio, and Ari Harju. "Force and heat current formulas for many-body potentials in molecular dynamics simulations with applications to thermal conductivity calculations." In: *Phys. Rev. B* 92 (9 2015), p. 094301. DOI: 10.1103/PhysRevB.92.094301.

[116] Zheyong Fan, Zezhu Zeng, Cunzhi Zhang, Yanzhou Wang, Keke Song, Haikuan Dong, Yue Chen, and Tapio Ala-Nissila. "Neuroevolution machine learning potentials: Combining high accuracy and low cost in atomistic simulations and application to heat transport." In: *Phys. Rev. B* 104 (10 2021), p. 104309. DOI: 10.1103/PhysRevB.104.104309.

[117] M. Bonetti, S. Nakamae, M. Roger, and P. Guenoun. "Huge Seebeck coefficients in nonaqueous electrolytes." In: *The Journal of Chemical Physics* 134.11 (2011), p. 114513. DOI: 10.1063/1.3561735.

[118] Hui Wang, Dan Zhao, Zia Ulla Khan, Skomantas Puzinas, Magnus P. Jonsson, Magnus Berggren, and Xavier Crispin. "Ionic Thermoelectric Figure of Merit for Charging of Supercapacitors." In: *Advanced Electronic Materials* 3.4 (2017), p. 1700013. DOI: https://doi.org/10.1002/aelm.201700013.

[119] Bruce J. Palmer. "Transverse-current autocorrelation-function calculations of the shear viscosity for molecular liquids." In: *Phys. Rev. E* 49 (1 1994), pp. 359–366. DOI: 10.1103/PhysRevE.49.359.

[120] Bruce J. Palmer. "Calculation of thermal-diffusion coefficients from plane-wave fluctuations in the heat energy density." In: *Phys. Rev. E* 49 (3 1994), pp. 2049–2057. DOI: 10.1103/PhysRevE.49.2049.

[121] Bingqing Cheng and Daan Frenkel. "Computing the Heat Conductivity of Fluids from Density Fluctuations." In: *Phys. Rev. Lett.* 125 (13 2020), p. 130602. DOI: 10.1103/PhysRevLett.125.130602.

[122] Enrico Drigo, Maria Grazia Izzo, and Stefano Baroni. "Heat conductivity from energy-density fluctuations." In: *The Journal of Chemical Physics* 159.18 (2023), p. 184107. ISSN: 0021-9606. DOI: 10.1063/5.0168732.

[123] L.A Rowley, D Nicholson, and N.G Parsonage. "Monte Carlo grand canonical ensemble calculation in a gas-liquid transition region for 12-6 Argon." In: *J. Comput. Phys.* 17.4 (1975), pp. 401–414. ISSN: 0021-9991. DOI: https://doi.org/10.1016/0021-9991(75)90042-X.

[124] B. W. H. van Beest, G. J. Kramer, and R. A. van Santen. "Force fields for silicas and aluminophosphates based on ab initio calculations." In: *Phys. Rev. Lett.* 64 (16 1990), pp. 1955–1958. DOI: 10.1103/PhysRevLett.64.1955.

[125] Alfredo Fiorentino and Stefano Baroni. "From Green-Kubo to the full Boltzmann kinetic approach to heat transport in crystals and glasses." In: *Phys. Rev. B* 107 (5 2023), p. 054311. DOI: 10.1103/PhysRevB.107.054311.

[126] Leyla Isaeva, Giuseppe Barbalinardo, Davide Donadio, and Stefano Baroni. "Modeling heat transport in crystals and glasses from a unified lattice-dynamical approach." In: *Nature Communications* 10.1 (2019), p. 3853. ISSN: 2041-1723. DOI: 10.1038/s41467-019-11572-4.

[127] Lorenzo Monacelli, Raffaello Bianco, Marco Cherubini, Matteo Calandra, Ion Errea, and Francesco Mauri. "The stochastic self-consistent harmonic approximation: calculating vibrational properties of materials with full quantum and anharmonic effects." In: *Journal of Physics: Condensed Matter* 33.36 (2021), p. 363001. DOI: 10.1088/1361-648X/ac066b. URL: https://dx.doi.org/10.1088/1361-648X/ac066b.

[128] Giuseppe Barbalinardo, Zekun Chen, Nicholas W. Lundgren, and Davide Donadio. "Efficient anharmonic lattice dynamics calculations of thermal transport in crystalline and disordered solids." In: *Journal of Applied Physics* 128.13 (2020), p. 135104. ISSN: 0021-8979. DOI: 10.1063/5.0020443.

[129] Philip B. Allen and Joseph L. Feldman. "Thermal conductivity of disordered harmonic solids." In: *Phys. Rev. B* 48 (17 1993), pp. 12581–12588. DOI: 10.1103/PhysRevB.48.12581.

[130] Philip B Allen and Jonathan Kelner. "Evolution of a vibrational wave packet on a disordered chain." In: *Am. J. Phys.* 66.6 (1998), pp. 497–506. DOI: 10.1119/1.18890.

[131] Philip B Allen and Joseph JL Feldman. "Thermal Conductivity of Glasses: Theory and Application to Amorphous Si." In: *Phys. Rev. Lett.* 62.6 (1989), pp. 645–648. DOI: 10.1103/PhysRevLett.62.645.

[132] Joseph L. Feldman, Philip B. Allen, and Scott R. Bickham. "Numerical study of low-frequency vibrations in amorphous silicon." In: *Phys. Rev. B* 59 (5 1999), pp. 3551–3559. DOI: 10.1103/PhysRevB.59.3551.

[133] Jaroslav Fabian Philip B. Allen Joseph L. Feldman and Frederick Wooten. "Diffusons, locons and propagons: Character of atomic vibrations in amorphous Si." In: *Philosophical Magazine B* 79.11-12 (1999), pp. 1715–1731. DOI: 10.1080/13642819908223054.

[134] R. Peierls. "Zur kinetischen Theorie der Wärmeleitung in Kristallen." In: *Annalen der Physik* 395.8 (1929), pp. 1055–1101. DOI: https://doi.org/10.1002/andp.19293950803.

[135] A. Ward, D. A. Broido, Derek A. Stewart, and G. Deinzer. "Ab initio theory of the lattice thermal conductivity in diamond." In: *Phys. Rev. B* 80 (12 2009), p. 125203. DOI: 10.1103/PhysRevB.80.125203.

[136] Andrea Cepellotti and Nicola Marzari. "Thermal Transport in Crystals as a Kinetic Theory of Relaxons." In: *Phys. Rev. X* 6 (4 2016), p. 041013. DOI: 10.1103/PhysRevX.6.041013.

[137] Michele Simoncelli, Nicola Marzari, and Francesco Mauri. "Unified theory of thermal transport in crystals and glasses." In: *Nature Physics* 15.8 (2019), pp. 809–813. ISSN: 1745-2481. DOI: 10.1038/s41567-019-0520-x.

[138] Michele Simoncelli, Nicola Marzari, and Francesco Mauri. "Wigner Formulation of Thermal Transport in Solids." In: *Phys. Rev. X* 12 (4 2022), p. 041011. DOI: 10.1103/PhysRevX.12.041011.

[139] Anees Pazhedath, Lorenzo Bastonero, Nicola Marzari, and Michele Simoncelli. "First-principles characterization of thermal conductivity in $LaPO_4$-based alloys." In: *Phys. Rev. Appl.* 22 (2 2024), p. 024064. DOI: 10.1103/PhysRevApplied.22.024064.

[140] Enrico Di Lucente, Michele Simoncelli, and Nicola Marzari. "Crossover from Boltzmann to Wigner thermal transport in thermoelectric skutterudites." In: *Phys. Rev. Res.* 5 (3 2023), p. 033125. DOI: 10.1103/PhysRevResearch.5.033125.

[141] Michele Simoncelli, Francesco Mauri, and Nicola Marzari. "Thermal conductivity of glasses: first-principles theory and applications." In: *npj Computational Materials* 9.1 (2023), p. 106. ISSN: 2057-3960. DOI: 10.1038/s41524-023-01033-4.

[142] Alfredo Fiorentino, Paolo Pegolo, and Stefano Baroni. "Hydrodynamic finite-size scaling of the thermal conductivity in glasses." In: *npj Computational Materials* 9.1 (2023), p. 157. ISSN: 2057-3960. DOI: 10.1038/s41524-023-01116-2.

[143] W. Schirmacher. "Thermal conductivity of glassy materials and the "boson peak"." In: *Europhysics Letters* 73.6 (2006), p. 892. DOI: 10.1209/epl/i2005-10471-9. URL: https://dx.doi.org/10.1209/epl/i2005-10471-9.

[144] Jason M. Larkin and Alan J. H. McGaughey. "Thermal conductivity accumulation in amorphous silica and amorphous silicon." In: *Phys. Rev. B* 89 (14 2014), p. 144303. DOI: 10.1103/PhysRevB.89.144303.

[145] Alfredo Fiorentino, Enrico Drigo, Stefano Baroni, and Paolo Pegolo. "Unearthing the foundational role of anharmonicity in heat transport in glasses." In: *Phys. Rev. B* 109 (22 2024), p. 224202. DOI: 10.1103/PhysRevB.109.224202.

[146] A. F. Ioffe and A. R. Regel. In: *Progresses in Semiconductors*. Ed. by A. F. Gibson. Vol. 4. Heywood, London, 1960, p. 237.

[147] Philip B. Allen, Joseph L. Feldman, Jaroslav Fabian, and Frederick Wooten. "Diffusons, locons and propagons: Character of atomic vibrations in amorphous Si." In: *Phil. Mag. B* 79.11-12 (1999), pp. 1715–1731. DOI: 10.1080/13642819908223054.

[148] Allan Griffin. "Brillouin Light Scattering from Crystals in the Hydrodynamic Region." In: *Rev. Mod. Phys.* 40 (1 1968), pp. 167–205. DOI: 10.1103/RevModPhys.40.167.

[149] Maria G. Izzo, Giancarlo Ruocco, and Stefano Cazzato. "The Mixing of Polarizations in the Acoustic Excitations of Disordered Media With Local Isotropy." In: *Frontiers in Physics* 6 (2018). ISSN: 2296-424X. DOI: 10.3389/fphy.2018.00108.

[150] M. G. Izzo, B. Wehinger, S. Cazzato, A. Matic, C. Masciovecchio, A. Gessini, and G. Ruocco. "Rayleigh scattering and disorder-induced mixing of polarizations in amorphous solids at the nanoscale: 1-octyl-3-methylimidazolium chloride glass." In: *Phys. Rev. B* 102 (21 2020), p. 214309. DOI: 10.1103/PhysRevB.102.214309.

[151] Jaeyun Moon, Raphaël P. Hermann, Michael E. Manley, Ahmet Alatas, Ayman H. Said, and Austin J. Minnich. "Thermal acoustic excitations with atomic-scale wavelengths in amorphous silicon." In: *Phys. Rev. Mater.* 3 (6 2019), p. 065601. DOI: 10.1103/PhysRevMaterials.3.065601.

[152] C Masciovecchio, G Baldi, S Caponi, L Comez, S Di Fonzo, D Fioretto, A Fontana, A Gessini, SC Santucci, F Sette, et al. "Evidence for a crossover in the frequency dependence of the acoustic attenuation in vitreous silica." In: *Physical review letters* 97.3 (2006), p. 035501.

[153] Giacomo Baldi, Valentina M Giordano, and Giulio Monaco. "Elastic anomalies at terahertz frequencies and excess density of vibrational states in silica glass." In: *Phys. Rev. B* 83.17 (2011), p. 174203. DOI: 10.1103/PhysRevB.83.174203.

[154] Jeffrey L. Braun, Christopher H. Baker, Ashutosh Giri, Mirza Elahi, Kateryna Artyushkova, Thomas E. Beechem, Pamela M. Norris, Zayd C. Leseman, John T. Gaskins, and Patrick E. Hopkins. "Size effects on the thermal conductivity of amorphous silicon thin films." In: *Phys. Rev. B* 93 (14 2016), p. 140201. DOI: 10.1103/PhysRevB.93.140201.

[155] Volker L. Deringer, Noam Bernstein, Albert P. Bartók, Matthew J. Cliffe, Rachel N. Kerber, Lauren E. Marbella, Clare P. Grey, Stephen R. Elliott, and Gábor Csányi. "Realistic Atomistic Structure of Amorphous Silicon from Machine-Learning-Driven Molecular Dynamics." In: *J. Phys. Chem. Lett.* 9.11 (2018), pp. 2879–2885. DOI: 10.1021/acs.jpclett.8b00902.

[156] Yuping He, Davide Donadio, and Giulia Galli. "Heat transport in amorphous silicon: Interplay between morphology and disorder." In: *Appl. Phys. Lett.* 98.14 (2011), p. 144101. ISSN: 0003-6951. DOI: 10.1063/1.3574366.

[157] B. L. Zink, R. Pietri, and F. Hellman. "Thermal Conductivity and Specific Heat of Thin-Film Amorphous Silicon." In: *Phys. Rev. Lett.* 96 (5 2006), p. 055902. DOI: 10.1103/PhysRevLett.96.055902.

[158] L Wieczorek, HJ Goldsmid, and GL Paul. "Thermal conductivity of amorphous films." In: *Thermal Conductivity 20* (1989), pp. 235–241.

[159] Tsuneyuki Yamane, Naoto Nagai, Shin-ichiro Katayama, and Minoru Todoki. "Measurement of thermal conductivity of silicon dioxide thin films using a 3ω method." In: *J. Appl. Phys.* 91.12 (2002), pp. 9772–9776. ISSN: 0021-8979. DOI: 10.1063/1.1481958.

[160] Xiao Liu, J. L. Feldman, D. G. Cahill, R. S. Crandall, N. Bernstein, D. M. Photiadis, M. J. Mehl, and D. A. Papaconstantopoulos. "High Thermal Conductivity of a Hydrogenated Amorphous Silicon Film." In: *Phys. Rev. Lett.* 102 (3 2009), p. 035901. DOI: 10.1103/PhysRevLett.102.035901.

[161] Ho-Soon Yang, David G. Cahill, X. Liu, J. L. Feldman, R. S. Crandall, B. A. Sperling, and J. R. Abelson. "Anomalously high thermal conductivity of amorphous Si deposited by hot-wire chemical vapor deposition." In: *Phys. Rev. B* 81 (10 2010), p. 104203. DOI: 10.1103/PhysRevB.81.104203.

[162] David G. Cahill, M. Katiyar, and J. R. Abelson. "Thermal conductivity of a-Si:H thin films." In: *Phys. Rev. B* 50 (9 1994), pp. 6077–6081. DOI: 10.1103/PhysRevB.50.6077.

[163] ET Swartz and RO Pohl. "Thermal resistance at interfaces." In: *Appl. Phys. Lett.* 51.26 (1987), pp. 2200–2202.

[164] Nathalie Vast and Stefano Baroni. "Effects of isotopic disorder on the Raman spectra of crystals: Theory and ab initio calculations for diamond and germanium." In: *Phys. Rev. B* 61 (14 2000), pp. 9387–9392. DOI: 10.1103/PhysRevB.61.9387.

[165] C.C. Paige. "Accuracy and effectiveness of the Lanczos algorithm for the symmetric eigenproblem." In: *Linear Algebra Appl.* 34 (1980), pp. 235–258. DOI: 10.1016/0024-3795(80)90167-6.

[166] J. Tersoff. "Empirical interatomic potential for silicon with improved elastic properties." In: *Phys. Rev. B* 38 (14 1988), pp. 9902–9905. DOI: 10.1103/PhysRevB.38.9902.

[167] Priya Vashishta, Rajiv K. Kalia, Aiichiro Nakano, and José Pedro Rino. "Interaction potential for silicon carbide: A molecular dynamics study of elastic constants and vibrational density of states for crystalline and amorphous silicon carbide." In: *J. Appl. Phys.* 101.10 (2007), p. 103515. ISSN: 0021-8979. DOI: 10.1063/1.2724570.

[168] José P. Rino, Ingvar Ebbsjö, Paulo S. Branicio, Rajiv K. Kalia, Aiichiro Nakano, Fuyuki Shimojo, and Priya Vashishta. "Short- and intermediate-range structural correlations in amorphous silicon carbide: A molecular dynamics study." In: *Phys. Rev. B* 70 (4 2004), p. 045207. DOI: 10.1103/PhysRevB.70.045207.

[169] S.-M. Lee and David G. Cahill. "Heat transport in thin dielectric films." In: *Journal of Applied Physics* 81.6 (1997), pp. 2590–2595. ISSN: 0021-8979. DOI: 10.1063/1.363923.

[170] Augustus Matthiessen and Moritz von Bose. "On the influence of temperature on the electric conducting power of metals." In: *Philos. Trans. R. Soc.* 152 (1862), pp. 1–27. DOI: 10.1098/rstl.1862.0001.

[171] O. V. Dolgov, D. A. Kirzhnits, and E. G. Maksimov. "On an admissible sign of the static dielectric function of matter." In: *Rev. Mod. Phys.* 53 (1 1981), pp. 81–93. DOI: 10.1103/RevModPhys.53.81.

[172] Philippe A. Bopp, Alexei A. Kornyshev, and Godehard Sutmann. "Static Nonlocal Dielectric Function of Liquid Water." In: *Phys. Rev. Lett.* 76 (8 1996), pp. 1280–1283. DOI: 10.1103/PhysRevLett.76.1280.

[173] Jean-Pierre Hansen and Ian R. McDonald. "Chapter 10 - Ionic Liquids." In: *Theory of Simple Liquids (Third Edition)*. Ed. by Jean-Pierre Hansen and Ian R. McDonald. Third Edition. Burlington: Academic Press, 2006, pp. 291–340. ISBN: 978-0-12-370535-8. DOI: https://doi.org/10.1016/B978-012370535-8/50012-4.

[174] O. V. Dolgov and E. G. Maksimov. "The sign of the static dielectric constant of simple metals." In: *JETP Letters* 28 (1 1978), p. 3. URL: http://jetpletters.ru/ps/0/article_23810.shtml.

[175] D A Kirzhnits. "Are the Kramers-Kronig relations for the dielectric permittivity of a material always valid?" In: *Soviet Physics Uspekhi* 19.6 (1976), p. 530. DOI: 10.1070/PU1976v019n06ABEH005268.

[176] V D Gorobchenko and Evgenii G Maksimov. "The dielectric constant of an interacting electron gas." In: *Soviet Physics Uspekhi* 23.1 (1980), p. 35. DOI: 10.1070/PU1980v023n01ABEH004860.

[177] David Pines. ""Extended Electron-Gas Hamiltonian" — an Author's Comment." In: *Phys. Rev. B* 2 (5 1970), pp. 1424–1425. DOI: 10.1103/PhysRevB.2.1424.

[178] Saverio Moroni, David M. Ceperley, and Gaetano Senatore. "Static Response and Local Field Factor of the Electron Gas." In: *Phys. Rev. Lett.* 75 (4 1995), pp. 689–692. DOI: 10.1103/PhysRevLett.75.689.

[179] Philippe A. Bopp, Alexei A. Kornyshev, and Godehard Sutmann. "Frequency and wave-vector dependent dielectric function of water: Collective modes and relaxation spectra." In: *The Journal of Chemical Physics* 109.5 (1998), pp. 1939–1958. ISSN: 0021-9606. DOI: 10.1063/1.476884.

[180] A. Fasolino, M. Parrinello, and M.P. Tosi. "Static dielectric behavior of charged fluids near freezing." In: *Physics Letters A* 66.2 (1978), pp. 119–121. ISSN: 0375-9601. DOI: https://doi.org/10.1016/0375-9601(78)90013-0.

[181] G. Monet, F. Bresme, A. Kornyshev, and H. Berthoumieux. "Nonlocal Dielectric Response of Water in Nanoconfinement." In: *Phys. Rev. Lett.* 126 (21 2021), p. 216001. DOI: 10.1103/PhysRevLett.126.216001.

[182] P. Bopp, G. Jancsó, and K. Heinzinger. "An improved potential for non-rigid water molecules in the liquid phase." In: *Chemical Physics Letters* 98.2 (1983), pp. 129–133. ISSN: 0009-2614. DOI: https://doi.org/10.1016/0009-2614(83)87112-7.

[183] A.G. Kalinichev and K. Heinzinger. "Molecular dynamics of supercritical water: A computer simulation of vibrational spectra with the flexible BJH potential." In: *Geochimica et Cosmochimica Acta* 59.4 (1995), pp. 641–650. ISSN: 0016-7037. DOI: https://doi.org/10.1016/0016-7037(94)00289-X.

[184] Imre Ruff and Dennis J. Diestler. "Isothermal–isobaric molecular dynamics simulation of liquid water." In: *The Journal of Chemical Physics* 93.3 (1990), pp. 2032–2042. ISSN: 0021-9606. DOI: 10.1063/1.459080.

[185] Bingqing Cheng. *Latent Ewald summation for machine learning of long-range interactions*. 2024. arXiv: 2408.15165 [cs.LG]. URL: https://arxiv.org/abs/2408.15165.

[186] Stephen J. Cox. "Dielectric response with short-ranged electrostatics." In: *Proceedings of the National Academy of Sciences* 117.33 (2020), pp. 19746–19752. DOI: 10.1073/pnas.2005847117.

[187] Ang Gao and Richard C. Remsing. "Self-consistent determination of long-range electrostatics in neural network potentials." In: *Nature Communications* 13.1 (2022), p. 1572. ISSN: 2041-1723. DOI: 10.1038/s41467-022-29243-2.

[188] Jean Pierre Hansen and Ian R. McDonald. "Statistical mechanics of dense ionized matter. IV. Density and charge fluctuations in a simple molten salt." In: *Phys. Rev. A* 11 (6 1975), pp. 2111–2123. DOI: 10.1103/PhysRevA.11.2111.

[189] John G. Kirkwood and Frank P. Buff. "The Statistical Mechanical Theory of Solutions. I." In: *The Journal of Chemical Physics* 19.6 (1951), pp. 774–777. ISSN: 0021-9606. DOI: 10.1063/1.1748352.

[190] Albert P. Bartók, Mike C. Payne, Risi Kondor, and Gábor Csányi. "Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons." In: *Phys. Rev. Lett.* 104 (13 2010), p. 136403. DOI: 10.1103/PhysRevLett.104.136403.

[191] Yuzhi Zhang, Haidi Wang, Weijie Chen, Jinzhe Zeng, Linfeng Zhang, Han Wang, and E Weinan. "DP-GEN: A concurrent learning platform for the generation of reliable deep learning based potential energy models." In: *Comp. Phys. Commun.* 253 (2020), p. 107206. DOI: 10.1016/j.cpc.2020.107206.

[192] John P. Perdew, Adrienn Ruzsinszky, Gábor I. Csonka, Oleg A. Vydrov, Gustavo E. Scuseria, Lucian A. Constantin, Xiaolan Zhou, and Kieron Burke. "Restoring the Density-Gradient Expansion for Exchange in Solids and Surfaces." In: *Phys. Rev. Lett.* 100 (13 2008), p. 136406. DOI: 10.1103/PhysRevLett.100.136406.

[193] Paolo Giannozzi, Stefano Baroni, Nicola Bonini, Matteo Calandra, Roberto Car, Carlo Cavazzoni, Davide Ceresoli, Guido L Chiarotti, Matteo Cococcioni, Ismaila Dabo, et al. "QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials." In: *J. Condens. Matter Phys.* 21.39 (2009), p. 395502. DOI: 10.1088/0953-8984/21/39/395502.

[194] Paolo Giannozzi, Oliviero Andreussi, Thomas Brumme, Oana Bunau, M Buongiorno Nardelli, Matteo Calandra, Roberto Car, Carlo Cavazzoni, Davide Ceresoli, Matteo Cococcioni, et al. "Advanced capabilities for materials modelling with Quantum ESPRESSO." In: *J. Condens. Matter Phys.* 29.46 (2017), p. 465901. DOI: 10.1088/1361-648X/aa8f79.

[195] Paolo Giannozzi, Oscar Baseggio, Pietro Bonfà, Davide Brunato, Roberto Car, Ivan Carnimeo, Carlo Cavazzoni, Stefano De Gironcoli, Pietro Delugas, Fabrizio Ferrari Ruffino, et al. "Quantum ESPRESSO toward the exascale." In: *J. Chem. Phys.* 152.15 (2020). DOI: 10.1063/5.0005082.

[196] Ivan Carnimeo, Fabio Affinito, Stefano Baroni, Oscar Baseggio, Laura Bellentani, Riccardo Bertossa, Pietro Davide Delugas, Fabrizio Ferrari Ruffino, Sergio Orlandini, Filippo Spiga, et al. "Quantum ESPRESSO: One further step toward the exascale." In: *Journal of Chemical Theory and Computation* 19.20 (2023), pp. 6992–7006. DOI: 10.1021/acs.jctc.3c00249.

[197] P. E. Blöchl. "Projector augmented-wave method." In: *Phys. Rev. B* 50 (24 1994), pp. 17953–17979. DOI: 10.1103/PhysRevB.50.17953.

[198] Nicola Marzari, David Vanderbilt, Alessandro De Vita, and M. C. Payne. "Thermal Contraction and Disordering of the Al(110) Surface." In: *Phys. Rev. Lett.* 82 (16 1999), pp. 3296–3299. DOI: 10.1103/PhysRevLett.82.3296.