GRADO EN QUÍMICA

Trabajo Fin de Grado

Título: Análisis de la conductividad térmica de la perovskita BaSnS3

Autor: Francisco Morales Martínez

Sevilla, 2023

Quiero dedicarle este trabajo a mi padre por enseñarme desde pequeño, la ilusión de descubrir, de aprender y tener una mentalidad matemática. A mi madre, por enseñarme a tener un gran corazón y que soy capaz de lo que me proponga si le echo ganas. A mi pareja, por estar en esta etapa tan difícil, apoyándome en las caídas y celebrando las victorias. A todas las personas que he conocido en mi vida, porque como bien dijo el filósofo José Ortega y Gasset : "Yo soy yo y mi circunstancia". Inclusive gracias a usted lector, que está siendo parte de esta circunstancia.

Índice

Resumen **3**

1 Introduction **5**

1.1 Thermoelectric Material 5

1.1.1 Seebeck Effect 6

1.1.2 Peltier Effect 7

1.2 Figure of merit in Thermoelectric 8

1.3 Perovskites 9

2 Objetivos **11**

3 Metodología **13**

3.1 Teoría del Funcional de la Densidad 13

3.2 La ecuación del transporte de Boltzmann 14

3.3 Cálculo de las constantes de fuerza 16

3.4 Aprendizaje automático y Machine Learning 18

3.5 Detalles Computacionales 19

4 Resultados y discusión **21**

4.1 Estructura 21

4.2 Validación de la metodología 21

4.3 Conductividad térmica del $BaSnS_3$ 23

5 Conclusions **29**

Bibliografía **33**

Resumen

Las perovskitas de calcogenuro, están siendo estudiadas por sus propiedades termoeléctricas, al ser materiales termoeléctricos de bajo costo. Para determinar si un material tiene una buena eficiencia como termoeléctrico, se utiliza la llamada figura de mérito (ZT), que depende de varios factores. En particular, es inversamente proporcional a la conductividad térmica de la red. El cálculo de la conductividad térmica de la red se realiza mediante la solución para la ecuación de transporte de Boltzman para fonones (BTE). Para ello son necesarios un conjunto de constantes de fuerza de segundo y tercer orden. Estas se obtendrán mediante una metodología de aprendizaje automático, a partir del cálculo de las fuerzas sobre los átomos en estructuras distorsionadas empleando la teoría del funcional de densidad (DFT).

En este trabajo fin de grado se van a calcular la conductividad térmica de la red de la perovskita $BaSnS_3$ y se analizarán los factores que relacionan la estructura del material y su composición con dicha magnitud. La perovskita $BaSnS_3$ tiene dos estructuras posibles, ambas de tipo ortorrómbico, pero una de ellas tiene una celda unidad de forma alargada, y se la denomina "needle-like", siendo la más estable. La otra estructura, que no ha sido estudiada previamente, se va a mostrar que es dinámicamente estable y que, por tanto, merece la pena ser investigada.

En este trabajo se ha utilizado métodos de aprendizaje automáticos para sacar las constantes de fuerzas de segundo y tercer orden del $BaSnS_3$. Con esto se obtienen los valores de las velocidades de grupo y los tiempos de vida medio de los fonones, los cuales son variables de las que depende la conductividad térmica. Al variar κ_l con la temperatura, se realizó a varias temperaturas para ver su tendencia. Por último, se examinó la dependencia de κ_l con la microestructura del material.

1 Introduction

The use of energy is increasingly in demand, either due to the use of devices that consume more or due to the increase of the world population, specially in developing countries. [1] To satisfy this demand, society combines different energy resources, from fossil fuel to renewable energies. However, we are facing a problem that goes beyond the energy sources. In all cases, it is observed that not all of the energy produced is used, most of it is wasted as heat. That is why we have to focus on improve the efficiency of our energy grid. To tackle this problem, thermoelectric materials has been proposed as an ideal solution, recycling part of the wasted heat as electricity. they began to investigate and go deeper into thermoelectric materials, which can make the energy yield much higher, and with this reduce the amount of consumption and thus the demand. Since the beginning of the use of thermoelectric materials both at an industrial level and at a user level, the reduction in energy comsuption has decreased. For this reason, its production and research is booming, their use is extending to wider waste-heat-recovery technologies with the overall thermoelectric generator market to grow from $406 M in 2021 to $635 M by 2026. [2]

1.1 Thermoelectric Material

Thermoelectricity, the field concerned with the conversion of heat energy into electrical energy and vice versa, revolves around the use of thermoelectric materials that possess unique properties enabling them to generate electricity when exposed to a temperature gradient or produce a cooling effect when an electric current is applied. The thermoelectric effect, comprising the Seebeck effect, Peltier effect, and Thomson effect, underlies the conversion of thermal energy into electrical energy and the reverse process of using electrical energy for cooling purposes. [3] The efficiency of energy conversion in thermoelectric materials is quantified by the thermoelectric figure of merit, often denoted as ZT. This dimensionless value incorporates material properties such as the Seebeck coefficient, electrical conductivity, and thermal conductivity. A higher ZT value indicates a more efficient thermoelectric material, yielding larger electrical output in relation to the temperature difference and minimized heat losses. The potential applications of thermoelectric materials are vast. They hold promise for waste heat recovery in industrial processes, automotive

exhaust systems, and power generation from renewable energy sources. Thermoelectric devices also find utility in cooling electronics, space applications, and portable refrigeration. [4] Ongoing research focuses on the quest for efficient thermoelectric materials and the enhancement of their properties. Scientists and engineers explore new compounds, nanostructured materials, and advanced fabrication techniques in the pursuit of novel materials with improved thermoelectric performance. [5]

1.1.1 Seebeck Effect

The Seebeck effect is a fundamental principle in thermoelectricity that describes the generation of an electric voltage in a conductor or thermoelectric material when there is a temperature difference along its length. Discovered by physicist Thomas Johann Seebeck in 1821, this effect occurs due to the differential heating of charge carriers within the material. In metals, the charge carriers are primarily electrons, while in semiconductors, both electrons and holes contribute to charge transport. [3] When a temperature gradient is applied across the material, the higher-temperature region experiences greater thermal energy, increasing the kinetic energy of the charge carriers. This leads to the diffusion of charge carriers from the hotter region to the colder region, resulting in an accumulation of charges and the establishment of an electric potential difference or voltage across the material. The voltage generated is directly proportional to the applied temperature difference and is quantified by the Seebeck coefficient or thermoelectric power. Seebeck's of two junctions at different temperatures T_H and T_C is:

$$V = -S\left(T_H - T_C\right) \tag{1}$$

The Seebeck coefficient represents the ratio of the induced voltage to the temperature gradient and is expressed in units of volts per Kelvin (V/K) or microvolts per degree Celsius (μV/°C). This effect allows for the conversion of heat energy into electrical energy by utilizing temperature gradients, making thermoelectric materials attractive for applications such as waste heat recovery, power generation in limited spaces, and portable energy harvesting. [6] To maximize the efficiency of thermoelectric devices based on the Seebeck effect, researchers focus on optimizing material properties. This includes enhancing the Seebeck coefficient, increasing electrical conductivity to facilitate charge transport, and reducing thermal conductivity to minimize heat dissipation.

1.1.2 Peltier Effect

The Peltier effect, named after the renowned physicist Jean Charles Athanase Peltier, is a fascinating phenomenon within the field of thermoelectricity. It involves the interaction between heat and electric current at the junction of two dissimilar materials. This effect is fundamental to the functioning of thermoelectric devices, which have made significant contributions to energy conversion and thermal management in various industries and applications. [4, 7] Peltier's heat rate for the electric current I is:

$$dQ/dt = (\Pi_A - \Pi_B)\, I \tag{2}$$

$$\Pi_A = kT \ln(n_L) \quad \Pi_B = kT \ln(n_R) \tag{3}$$

where dQ/dt is the differential heat. Π_A and Π_B are the coefficients of each material. The coefficients depends, k is Boltzmann's constant, T is the temperature and $Ln(n)$ is the number of holes or excess electrons depending on whether it is type n or type p.

Consider a thermoelectric junction consisting of p-type and n-type semiconductors. When a direct electric current passes through this junction, the movement of electrons in the n-type material and the positively charged "holes" in the p-type material occurs in opposite directions. As a result, energy is transferred, leading to either the heating or cooling of the junction, depending on the direction of the current flow. At the starting point of the current, known as the "hot side" or "heating side," the thermoelectric junction absorbs heat from the surrounding environment, resulting in a cooling effect. This principle is utilized in thermoelectric refrigeration systems, where precise cooling of specific components or areas is required. Conversely, at the endpoint of the current, referred to as the "cold side" or "cooling side," the thermoelectric junction dissipates heat into the environment, leading to a heating effect. This concept is applied in thermoelectric heating systems and temperature control devices, where accurate regulation of temperature is essential. The Peltier effect is rooted in the intricate relationship between temperature differences and electric voltage or current. Semiconductors, with their unique ability to conduct both heat and electricity, play a critical role in this effect. The efficiency of the Peltier effect is quantified by the thermoelectric figure of merit (ZT), which considers properties such as the Seebeck coefficient, electrical conductivity, and thermal conductivity.

1.2 Figure of merit in Thermoelectric

The thermoelectric figure of merit, commonly referred to as ZT, is a crucial parameter in thermoelectric research. It provides a measure of the efficiency of a material in converting waste heat into usable electrical energy or vice versa. ZT is derived from several fundamental thermoelectric properties, including the Seebeck coefficient (S), electrical conductivity (α), and thermal conductivity (κ). The Seebeck coefficient represents the ability of a material to generate an electric voltage when subjected to a temperature gradient. It is a measure of the magnitude and polarity of the thermoelectric potential that arises across the material. A higher Seebeck coefficient indicates a larger voltage output and, thus, better thermoelectric performance. [8]

Electrical conductivity is a measure of a material's ability to conduct electric current. In the context of thermoelectric materials, a high electrical conductivity is desired to minimize power losses and maximize the efficiency of energy conversion. Materials with high carrier mobility and low resistivity tend to exhibit enhanced electrical conductivity. [8, 9]

Thermal conductivity, on the other hand, characterizes a material's ability to conduct heat. In thermoelectric applications, a low thermal conductivity is desired to minimize heat dissipation and enhance the temperature gradient across the material. This promotes efficient energy conversion by maximizing the temperature difference between the hot and cold sides of the device.

ZT combines these properties into a single dimensionless figure of merit, given by the equation:

$$ZT = \frac{S^2\sigma}{\kappa}T \tag{4}$$

where T is the temperature, σ is the electrical conductivity, κ is the thermal conductivity, and S is the Seebeck coefficient.

A higher ZT value indicates better thermoelectric performance, as it signifies a larger thermoelectric conversion efficiency. [10] The search for high ZT materials is motivated by the need for sustainable and efficient energy conversion technologies. Thermoelectric devices offer the potential to harness waste heat from various industrial processes and convert it into useful electrical power, thus contributing to energy conservation and reducing greenhouse gas emissions. Moreover, thermoelectric systems can also be utilized in refrigeration and cooling applications, offering advantages such as compactness, durability, and environmental friendliness.

By understanding the significance of ZT and conducting research to improve this figure of merit, scientists and engineers aim to develop novel thermoelectric materials with enhanced efficiency and performance. This involves exploring various strategies, including material design, nanostructuring, alloying, and optimization of carrier concentration, to manipulate the thermoelectric properties and achieve higher ZT values. [10, 11]

1.3 Perovskites

Perovskite materials have garnered significant attention in the field of Materials Science due to their remarkable properties and potential applications across various technological domains. The crystal structure known as perovskite, named after the mineral perovskite, follows a general formula ABX_3, where A and B represent cations and X denotes an anion. This structure consists of BX_6 octahedra that share corners, creating a three-dimensional network and enabling the exploration of a wide range of physical and chemical phenomena [12]. Among the various perovskite materials, $BaSnS_3$ is of particular interest as it belongs to the chalcogenide perovskite family. Chalcogenides encompass compounds containing elements from the chalcogen group such as sulfur (S), selenium (Se), or tellurium (Te). $BaSnS_3$ exhibits intriguing thermoelectric properties, making it a promising material for applications in energy conversion and waste heat recovery.

Understanding the structural characteristics of $BaSnS_3$ is crucial in elucidating its thermoelectric behavior. This involves comprehending the arrangement of atoms within the crystal lattice and the extent of deviation from the idealized perovskite structure. [13] These structural properties significantly influence the transport properties of $BaSnS_3$ and directly impact its thermoelectric performance.

To investigate the thermoelectric properties of $BaSnS_3$, it is essential to analyze the Boltzmann Transport Equation (BTE). The BTE provides a theoretical framework for describing transport phenomena in materials, including electrical conductivity, Seebeck coefficient, and thermal conductivity. By solving the BTE, it becomes possible to explore the contributions of electrons and phonons to the thermal and electrical transport in $BaSnS_3$ perovskite.

This work aims to examine the thermoelectric properties of $BaSnS_3$ perovskite through simulations and analysis utilizing the BTE. Moreover, we will employ machine learning techniques to enhance our understanding of the structure-property relationships and opti-

mize the thermoelectric performance of $BaSnS_3$ perovskite. By varying the cutoff and the number of cells, we will construct supercells to investigate the impact of different structural configurations on the thermoelectric behavior. Furthermore, the machine learning model will enable the prediction of the thermodynamic properties of $BaSnS_3$ perovskite across a wide temperature range, providing insights into its conductivity under different operating conditions. [14]

Through this research, we aspire to contribute to the advancement of thermoelectric materials and facilitate the design and development of efficient energy conversion devices. The integration of computational simulations, BTE analysis, and machine learning approaches offers a comprehensive and systematic methodology to investigate the thermoelectric properties of $BaSnS_3$ perovskite, offering valuable insights for future material design and optimization.

2 Objetivos

El objetivo principal de este estudio es calcular la conductividad térmica debido a las vibraciones de la red de la perovskita $BaSnS_3$ con el fin de explorar su potencial en dispositivos termoeléctricos. Para ello, se utilizarán cálculos basados en la teoría del funcional de la densidad, Machine Learning y la solución de la ecuación de transporte de Boltzmann para fonones. Para su consecución se han establecido distintos objeticos específicos.

- Comprender las bases de la teoría del funcional de la densidad (DFT)
- Entender los conceptos básicos y etapas principales en el proceso de elaboración de un modelo predictivo empleando algoritmos de Machine Learning.
- Entender los conceptos básicos del transporte de calor a través de las vibraciones de la red cristalina y el empleo de la solución de la ecuación de transporte de Boltzmman para fonones.
- Analizar las principales propiedades que afectan al transporte de calor

3 Metodología

3.1 Teoría del Funcional de la Densidad

Los métodos *ab initio*, desarrollados aplicando aproximaciones matemáticas a la resolución de la ecuación de Schrödinger, tienen el problema de la dependencia de la función de onda de las $3N$ coordenadas electrónicas. Esto hace que el coste computacional de estos métodos limite mucho su aplicación al estudio de sistemas de gran tamaño. Este problema no está presente en la Teoría del Funcional de la Densidad (DFT, por sus siglas en inglés), ya que su magnitud fundamental es la densidad electrónica, que depende sólo de tres variables. [15] Los métodos DFT se basan en los teoremas de Hohenberg-Kohn los cuales establecen que las propiedades en el estado fundamental de un sistema polielectrónico están determinadas de manera única por la densidad electrónica. Esto significa que si dos sistemas tienen la misma densidad electrónica, mostrarán la misma energía y propiedades. Las ecuaciones de Kohn-Sham introducen un conjunto de electrones ficticios no interactuantes que reproducen la misma densidad electrónica que el sistema real. [16–18]

Dentro de la DFT, la energía de un sistema polielectrónico se define como

$$E[\rho_0] = F_{HK}[\rho_0] + \int \rho_0(\mathbf{r}) V_{ext}(\mathbf{r}) d\mathbf{r} = T[\rho_0] + V_{ee}[\rho_0] + \int \rho_0(\mathbf{r}) V_{ext}(\mathbf{r}) d\mathbf{r} \quad (5)$$

donde $F_{HK}[\rho_0]$ es el funcional de Kohn-Sham, que contiene a los funcionales de energía cinética $T[\rho_0]$ y de repulsión electrón-electrón, $V_{ee}[\rho_0]$, y V_{ext} es el potencial externo al sistema electrónico (por ejemplo, el generado por los núcleos del sistema). La densidad electrónica ρ_0 que minimiza el valor de la energía $E[\rho_0]$ es la densidad del estado fundamental del sistema y la energía obtenida es la energía de dicho estado fundamental. Si bien la teoría DFT es exacta, el problema recae en la definición del operador $F_{HK}[\rho_0]$ en la ec. 5, para el que no se tiene una expresión exacta, salvo en casos muy sencillos.

Un avance fundamental para llegar a una aplicación práctica de la DFT fue la propuesta de Kohn y Sham de emplear un sistema ficticio de electrones no interactuantes que tuviera la misma densidad electrónica que el sistema real. [15, 19] De esta forma, la energía total se puede calcular ahora como

$$E[\rho_0] = T_S[\rho_0] + \int \rho_0(\mathbf{r}) V_{ext}(\mathbf{r}) d\mathbf{r} + \frac{1}{2} \int \frac{\rho_0(\mathbf{r_1}) \rho_0(\mathbf{r_2})}{|\mathbf{r_1} - \mathbf{r_2}|} d\mathbf{r_1} d\mathbf{r_2} + E_{XC}[\rho_0] \quad (6)$$

donde se reemplaza la energía cinética del sistema real, $T[\rho_0]$ por la del sistema ficticio, $T_S[\rho_0]$ y la repulsión electrónica real por el término $\frac{1}{2}\int \frac{\rho_0(\mathbf{r_1})\rho_0(\mathbf{r_2})}{|\mathbf{r_1}-\mathbf{r_2}|} d\mathbf{r_1} d\mathbf{r_2}$ que representa la interacción puramente culombiana de la nube electrónica consigo misma. El término $E_{XC}[\rho_0]$, denominado funcional de intercambio-correlación agrupa las correcciones que es necesario añadir (diferencia entre la energía cinética del sistema real y del sistema modelo, incorporación de la antisimetría de la función de onda y correlación electrónica) y para el que no se conoce una expresión explícita. [20] Los primeros trabajos que intentaron resolver el problema del funcional de intercambio-correlación se basaron en el modelo de gas de electrones y resultaron en la conocida como la aproximación LDA (LSDA en su versión con densidades electrónicas diferentes para los electrones de spin α y β). En sistemas moleculares, sin embargo, las aproximaciones realizadas por la LDA, fundamentalmente que la densidad electrónica varía de forma suave, dejan de ser aplicables ya que en la proximidades de los núcleos hay grandes variaciones de la densidad. Como consecuencia se desarrolló toda una nueva familia de funcionales (funcionales GGA) en los que se incluye una dependencia explícita del funcional de intercambio-correlación con el gradiente de la densidad electrónica, que incluye el funcional de Perdew-Burke-Ernzerhof (PBE) que se va a utilizar en este TFG. [21]

3.2 La ecuación del transporte de Boltzmann

La solución a la ecuación de transporte de Boltzmann para fonones, BTE, fue propuesta por Peierls por primera vez en 1929, y a día de hoy, se considera uno de los métodos más precisos para calcular la conductividad térmica. [22] Las primeras aproximaciones para su uso fueron establecidas por Callaway [23] y Allen [24], los cuales empleaban variables que se ajustaban a datos experimentales. Más tarde, la solución de la BTE fue combinada con la teoría de la perturbación del funcional de la densidad, DFT, por Deinzer para investigar los anchos de línea (tiempos de vida media) de fonones en semiconductores. La forma más precisa, fiable y sistemática de determinar la conductividad térmica de la red ha sido probablemente combinar la solución de la BTE con cálculos basados en la DFT. [25]

Existen distintas implementaciones desarrolladas en varios paquetes como, como AAPL, ShengBTE o Phono3py, para pronosticar las conductividades térmicas de la red de aislantes y semiconductores, sin embargo, todos ellos tienen en común el método diseñado por Omini en el cual la conductividad térmica es calculada de una manera

autoconsistente. [26]

$$\kappa_{\mathrm{l}}^{\alpha\beta} = \frac{1}{k_\beta T^2 \Omega N} \sum_{\lambda} f_0 \left(f_0 + 1\right) \left(h\omega_\lambda\right)^2 v_\lambda^\alpha F_\lambda^\beta \tag{7}$$

En esta ecuación los superíndices α y β son dos direcciones cartesianas y λ combina un vector de onda $\mathbf{q}$ y el índice del modo vibracional del fonón i, v_λ y ω_λ representan la velocidad de grupo y la frecuencia del modo vibracional λ, respectivamente. Además, $f^0(\omega_\lambda)$ representa la ocupación del estado vibracional λ a una cierta temperatura siguiendo una estadística de Bose-Einstein (los fonones son bosones). La conductividad se calcula utilizando una malla de puntos, $\mathbf{q}$, explorando el interior de la zona de Brillouin. El camino libre de F_λ se representa como:

$$F_\lambda = \tau_\lambda^0 \left(v_\lambda + \Delta_\lambda\right) \tag{8}$$

donde τ^λ es el tiempo de relajación del fonón y Δ_λ depende a su vez del camino libre medio, por lo que la BTE debe resolverse de manera iterativa. Si este segundo término es ignorado, nos encontramos en la denominada aproximación del tiempo de relajación (RTA) ampliamente utilizada. Las soluciones de RTA pronostican valores de κ_{l} que son menores que los pronosticados por el método autoconsistente. Por lo tanto, se requieren determinar tres parámetros importantes para resolver el BTE para fonones: τ^λ, v_λ y ω_λ.

Para obtener las frecuencias de vibración de cada modo vibracional es necesario construir y diagonalizar la matriz dinámica. Esta matriz dinámica es específica para cada punto **q** de la malla de puntos en la que se subdivide la zona de Brillouin. Para generar la matriz dinámica, son necesarias las masas de los átomos y las contantes de fuerza de segundo orden. La diagonalización de esta matriz da como resultado un conjunto de valores que denominamos autovalores y se corresponden a los cuadrados de las frecuencias de vibración de cada modo normal en ese punto **q** determinado. [27] De las derivadas de matrices dinámicas con respecto a los puntos **q**, es posible extraer la velocidad de grupo, v_λ, empleando el teorema de Hellmann-Feynman.

A la hora de calcular τ^λ es necesario tener en cuenta que existen destinos mecanismos por los cuales es posible que se produzca la dispersión ("scattering") del fonón. La presencia de elementos con distintos isótopos (y por lo consiguiente masas) hace que las frecuencias sean ligeramente distintas en función del isótopo, eso causa procesos de dispersión elásticos que disminuyen la conductividad térmica del material. Además, la presencia de defectos puntuales (dopantes, vacantes, ...), dislocaciones, o bordes de grano

son barreras que reducen el tiempo de vida media del los fonones reduciendo también conductividad térmica del material. En sistemas isotópicamene puros, estequiométricos y cristalinos, los procesos de dispersión comentados anteriormente no aparecen, sin embargo, sigue habiendo procesos de dispersión debido a la interacción entre fonones (interferencias entre ondas). Los procesos de dispersión más sencillos entre fonones son los de tercer orden y juegan un papel preponderante a temperaturas bajas y medias. Los procesos de interferencia entre más de tres fonones son poco probables y, por lo tanto, sólo son importantes a muy elevadas temperaturas.

Desde un punto de vista computacional, los procesos de dispersión entre fonones son los más costosos de calcular. Para introducirlos en la solución de la BTE es necesario calcular primero las contantes de fuerza de tercer orden. Los procesos de dispersión entre tres fonones pueden clasificarse de dos tipos: i) dos fonones interaccionan para dar un tercer fonón (proceso +), ii) un fonón da lugar a dos fonones (procesos -). En cualquiera de los casos es necesario que durante el proceso de dispersión se produzca una conservación del momento y de la energía. El alto coste computacional de los tiempos de vida media debido a los procesos de dispersión entre fonones reside en el gran número de constantes de fuerza de tercer orden con respecto a las de segundo (dos o tres órdenes de magnitud mayores en número). Con el fin de reducir el coste computacional es necesario aplicar dos estrategias fundamentales: i) análisis exhaustivo de la simetría del cristal para centrar todos los recursos en el cálculo exclusivo de las constantes de fuerza no equivalentes por simetría, y ii) la definición de “cutoffs” los cuales determinan distancias limites más allá de las cuales las contantes de fuerza se toman como cero.

3.3 Cálculo de las constantes de fuerza

La sección anterior pone de manifiesto la importancia del cálculo de las contantes de fuerza en la predicción de la conductividad térmica basándose en la solución de la BTE para fonones. Las constantes de fuerza de segundo son imprescindibles en la construcción de la matriz dinámica que a su vez se emplea en el cálculo de frecuencias y velocidades de grupo. Por otro lado, las constantes de fuerza de tercer orden son necesarias para el cálculo de los tiempos de vida media.

La manera tradicional de calcular las constantes de fuerza es a través del método de diferencias finitas. Para ello, se toma una celda unidad la cual ha sido optimizada con

anterioridad (el sistema se encuentra en el mínimo de la superficie de energía potencial) y se generan superceldas de mayor tamaño. A continuación, se estudia su simetría y se estudian las distorsiones no equivalentes. Para cada una de esas distorsiones se generarán superceldas en las cuales un único átomo es desplazado fuera de su posición de equilibrio en una dirección concreta. Al distorsionar la posición de un átomo, el sistema no se encuentra en el mínimo de la superficie de energía de potencial por lo que aparecen fuerzas sobre cada uno de los átomos de la celda. Esas fuerzas pueden ser calculadas utilizando paquetes basados en el DFT u otros métodos *ab-initio*. El cociente entre las fuerzas y la magnitud de la distorsión da como resultado la constante de fuerza de segundo orden. Un procedimiento similar se sigue para las constantes de fuerza de tercer orden, pero en este caso es necesario realizar distorsiones sobre dos átomos. Esta metodología es ampliamente utilizada y se encuentra en paquetes como Phonopy o ShengBTE.

Durante los últimos años, el cálculo de las constantes de fuerza a través de diferencias finitas está siendo remplazado por metodologías basadas en regresión y Machine Learning. Estas metodologías reducen entre uno y dos órdenes de magnitud el coste computacional. Al igual que en las diferencias finitas, es necesario optimizar la celda unidad y a continuación generar una supercelda. Sin embargo, en este caso en lugar de distorsionar uno o dos átomos en cada una de estas superceldas, se construye una distorsión generalizada de todos los átomos del sistema. A continuación, se calculan las fuerzas sobre cada uno de los átomos para poder generar un sistema de ecuaciones

$$\Psi_i^\alpha = -\sum_j \sum_\beta \Phi_{ij}^{\alpha\beta} r_j^\beta - \frac{1}{2} \sum_{j,k} \sum_{\beta,\gamma} \Phi_{ijk}^{\alpha\beta\gamma} r_j^\beta r_k^\gamma - \dots \tag{9}$$

donde Ψ_i^α es la fuerza sobre el átomo i en la dirección α, r_j^β representa la magnitud de la distorsión en la dirección β, r_k^γ representa la magnitud de la distorsión en la dirección γ y $\Phi_{ij}^{\alpha\beta}$ y $\Phi_{ijk}^{\alpha\beta\gamma}$ son las constantes de fuerza de segundo y tercer orden, respectivamente. [28] De este sistema se conocen tanto las fuerzas como las magnitudes de los desplazamientos por lo que es posible realizar una regresión numérica cuyos coeficientes serán las constantes de fuerza del sistema a estudio. Esta metodología es muy sensible al tipo y magnitud de las distorsiones realizadas por lo que es preciso un proceso de validación del modelo de regresión (ver detalles computaciones y resultados).

3.4 Aprendizaje automático y Machine Learning

Dentro de lo que llamamos Aprendizaje Automático podemos englobar una serie de métodos computacionales que lo que pretenden es convertir conjuntos de datos en modelos más allá de un tratamiento estadístico de los mismos. Mientras que un programa "tradicionalüsa una serie de datos de entrada para convertirlos, mediante una serie de reglas, en datos de salida, los algoritmos de Aprendizaje Automático lo que pretenden es inferir las reglas que relacionan unos datos de entrada con otros de salida. Estos datos se agrupan en tres categorías:

- **Datos de entrenamiento**: Se refiere a los elementos utilizados durante el proceso de aprendizaje. El algoritmo se encargará de analizar el conjunto de datos y determinar las combinaciones que generen el modelo predictivo más adecuado.
- **Datos de Validación**: Estos datos están relacionados con la predicción de las respuestas realizadas por el modelo final, con el objetivo de lograr un alto nivel de confianza en dichas predicciones.
- **Datos de prueba**: Son un conjunto de ejemplos utilizado para evaluar el rendimiento del modelo óptimo. Durante la evaluación, el modelo realiza predicciones de salida para los ejemplos del conjunto de prueba y los compara con los valores reales correspondientes a esos ejemplos.

Los modelos de aprendizaje automático se dividen en diferentes categorías dependiendo de cómo se realiza el aprendizaje que genera el modelo. En nuestro caso vamos a emplear lo que se llama aprendizaje supervisado, en el que los datos de entrenamiento son un conjunto de datos de entrada con sus correspondientes vectores objetivos. Es necesario comenzar por determinar de qué tipo van a ser los datos y seleccionar un conjunto de datos de entrenamiento. Este conjunto ha de ser representativo del uso que se le quiera dar posteriormente al modelo generado, suficientemente amplio y variado (para evitar generar un modelo que no se ajuste con los datos de validación, lo que indicaría un sobreentrenamiento del modelo). Si bien hay muchos algoritmos de selección de los datos de entrenamiento (LOO, SVM, EnRFE, . . .), muchos de ellos tienen un excesivo coste computacional. Como alternativas más eficientes tenemos la eliminación de características recursivas (Recursive Feature Elimination, RFE) y el método de mínimos cuadrados (Ordinary Least-Squares, OLS). [29]

El algoritmo RFE ofrece tiene un buen balance entre los resultados y el su coste computacional. Se basa en ir eliminando las funciones débiles (con poca participación en el modelo) y reentrenar el modelo. La eliminación de las funciones débiles se hace de una en una, para garantizar que no se va a degradar la precisión del modelo generado. [30, 31]

3.5 Detalles Computacionales

Estructura inicial. La estructura inicial del $BaSnS_3$ se tomó de la base de datos Materials Projects [32]

Cálculos DFT. La optimización de la celda unidad (posiciones de los átomos y parámetros de red) y el cálculo de fuerzas de las superceldas se ha realizado usando la teoría del funcional de la densidad (DFT) mediante el paquete VASP (Vienna Ab-initio Simulation Package). [33] Se ha empleado en funcional GGA desarrollado por Perdew, Burke y Ernzerhof (PBE) [21] para describir las energías y las fuerzas. Los electrones de valencia se han descrito empleando ondas planas con un *cutoff* de $500\,\mathrm{eV}$ y los electrones de core han sido descritos empleando pseudopotenciales optimizados para el funcional GGA descrito anteriormente. La estructura se ha considerado optimizada (mínimo de la superficie de energía potencial) cuando las fuerzas sobre todos los átomos eran menores a $1 \times 10^{-7}\,\mathrm{eV/\AA}$, empleando una malla de puntos de $6 \times 4 \times 6$ puntos **k** dentro de la zona de Brillouin. En cada paso de optimización y en los cálculos puntuales para el cálculo de fuerzas, la función de onda se ha dado por convergida cuando la diferencia de energía entro dos pasos consecutivos del cálculo autoconsistente era menor a $1 \times 10^{-9}\,\mathrm{eV}$.

Constantes de fuerza. Las constantes de fuerza han sido extraídas de un modelo realizado con una supercelda de tamaño $3 \times 2 \times 3$ en la cual los átomos han sido desplazados de sus posiciones de equilibrio. Las distorsiones realizadas se basan en los autovectores correspondientes a cada modo de vibración del material y su amplitud se ha adaptado a la que tendrían las vibraciones en el sistema a una temperatura de $300\,\mathrm{K}$. De esta manera se garantiza que i) los distorsiones tienen un sentido físico y ii) no es necesario incluir constantes de fuerza de orden mayor a 4 para poder conseguir una buena regresión. Primero, las fuerzas obtenidas se dividen en dos grupos, aproximadamente un 80 % se emplea en el entrenamiento (regresión) y el 20 % restante en la validación del modelo. El número de constantes de fuerza (coeficientes del modelo lineal) está determinado por los *cutoffs* de interacción que son optimizados (ver resultados). El número de superceldas empleadas

para el modelo, que a su vez determina el número de ecuaciones de nuestro sistema, se ha optimizado para reducir el error. Debido al gran número de constantes de fuerza y con con el fin de evitar fenómenos de sobreentrenamiento, se ha empleado la técnica de eliminación recursiva de características (RFE) que permite simplificar el modelo final y reducir el error en la etapa de validación. Finalmente, con el fin de asegurar que el modelo no es dependiente del método seguido en la subdivisión de datos entre entrenamiento y validación, se ha realizado una quíntuple validación cruzada. Todo el proceso descrito se ha realizado utilizando el paquete HiPhive. [28]

Cálculo de la conductividad térmica. La conductividad térmica se ha calculado empleando la solución de la ecuación de transporte de Boltzmann para fonones de manera iterativa empleando en código ShengBTE. [34] Los tiempos de vida media han sido calculados teniendo en cuenta tanto los procesos de dispersión entre tres fonones como los procesos de dispersión elásticos debido a la abundancia isotópica de cada uno de los elementos que constituyen el $BaSnS_3$. El consumo de memoria y tiempo de cálculo de este paso se ha optimizando utilizando un *scalebroad* de 0.05 y una malla de puntos $\mathbf{q}$ suficientemente densa que permite la convergencia de los valores de κ_l (ver sección de resultados)

4 Resultados y discusión

4.1 Estructura

La perovskitas de calcogenuro pueden presentar poliformismo. Además de la estructura cúbica, existen otros poliformos, pertenecientes a otros grupos espaciales, como ortorrómbico o hexagonal, en los cuales los octaedros se encuentran ligeramente rotados en la estructura. La estabilidad relativa de estos poliformos dependen de la composición química del material. Por ejemplo para $BaZrS_3$, se han reportado dos poliformos ortorrómbicos con el mismo grupo espacial. Para diferenciarlos, uno de ellos se le denomina poliformo con forma de aguja (needle-like) por la forma alargada de la celda, el otro es llamado simplemente ortorrómbico. Mientras que para el $BaZrS_3$, la fase más estable es la ortorrombica no alargada, para el $BaSnS_3$, la estructura "needle-like", es la reportada experimentalmente y de la cual Li *et al.* se han calculado sus propiedades tanto ópticas como de transporte electrónico y térmico. [12] A pesar de ser ligeramente menos estable que la estructura "needle-like"(alrededor de $0.075\,\mathrm{eV}$ por átomo), la estructura ortorrombica no alargada presenta estabilidad dinámica (ver más adelante en la sección de resultados), por lo que no se descarta la posibilidad de ser una fase metaestable y su potencial síntesis. Es por ello que nos centraremos en esta fase de aquí en adelante.

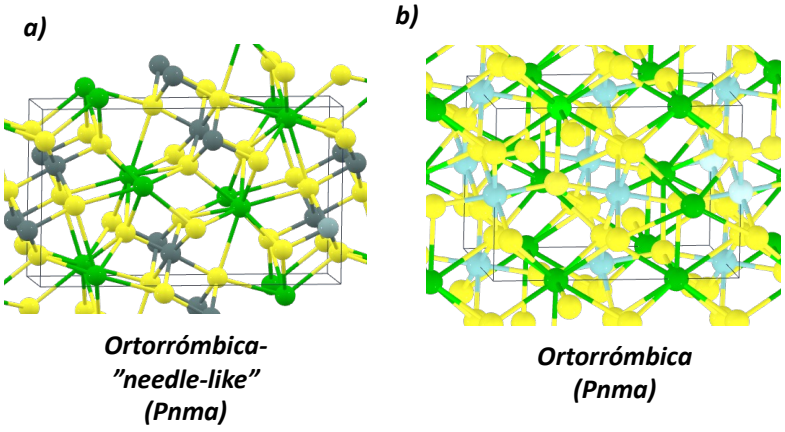

Figura 1: Comparación de las dos estructuras ortorrómbicas en las perovskitas de Ba.

4.2 Validación de la metodología

Antes de adentrarnos en el análisis de los resultados, es esencial llevar a cabo una minuciosa validación de la metodología empleada. Esta validación engloba la optimización

de múltiples parámetros, desde la elección del funcional utilizado en el cálculo de fuerzas hasta la adecuada selección de los puntos **q** empleados para evaluar la conductividad térmica. Asimismo, es crucial tener en cuenta las variables asociadas al entrenamiento del modelo, el cual proporciona las constantes de fuerza requeridas.

Resulta de vital importancia analizar el error cuadrático medio durante la obtención de las constantes de fuerza, tanto en la etapa de validación como en el entrenamiento del modelo. Para este proceso, se deben considerar tres variables: el número de superceldas utilizadas en el entrenamiento, el valor del radio de corte (cutoff) y el algoritmo empleado en la regresión. El objetivo principal consiste en lograr un modelo con el menor error en la validación, evitando la aparición de fenómenos indeseables como el sobreentrenamiento de datos (llamado overfitting). Este fenómeno consiste en el intento del algoritmo de forzar que la función se ajuste a todos los datos del entrenamiento, haciendo que aumente el error de la validación. Este análisis se ve reflejado en la Figura 2.

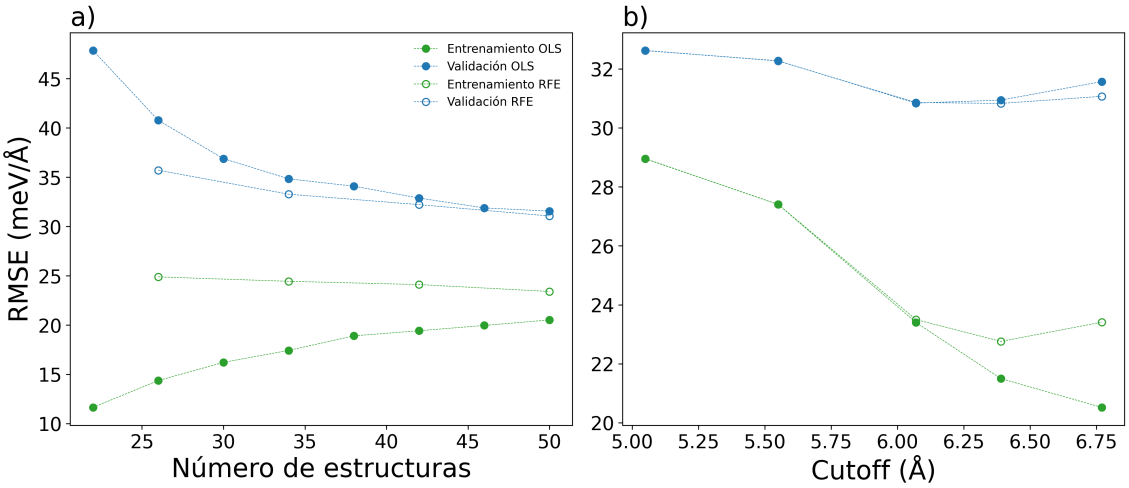

Figura 2: Comparación de OLS con RFE a diferentes nº de estructuras y cutoff

En primer lugar, se puede observar como en la Figura 2a), a medida que aumenta el número de datos (número de superceldas), el error cuadrático medio durante entrenamiento y validación tienden a converger. La principal diferencia entre OLS y RFE radica en la velocidad con la que RFE alcanza ese valor. Gracias a esto se puede determinar que el número óptimo de estructuras y con las cuales vamos a trabajar serán 50. Por otro lado, en la Figura 2b) se puede apreciar como a medida que se incrementa el cutoff y, por lo tanto, aumenta el número de variables o coeficientes, lo que puede llegar a causar un sobreenetrenamiento (aumento del RMSE de validación con el cutoff). Es relevante

destacar que la técnica RFE ayuda a mitigar ligeramente el sobreentrenamiento que ocurre con valores de cutoff más altos. Aunque esta reducción en el error de validación es mínima, se ha optado por utilizar RFE con un valor de cutoff de 6.4 Å.

Una vez validado el modelo generado con Machine Learning, es necesario validar la metodología empleada para el cálculo de la conductividad térmica. En este apartado es crítico definir una malla de puntos **q** dentro de la zona de Brillouin, que permita obtener un valor de conductividad térmica convergido. En la Figura 3) se puede observar como a un total de 16 puntos **q** el valor de κ_l converge a $0.74\,\mathrm{W\,m^{-1}\,K^{-1}}$.

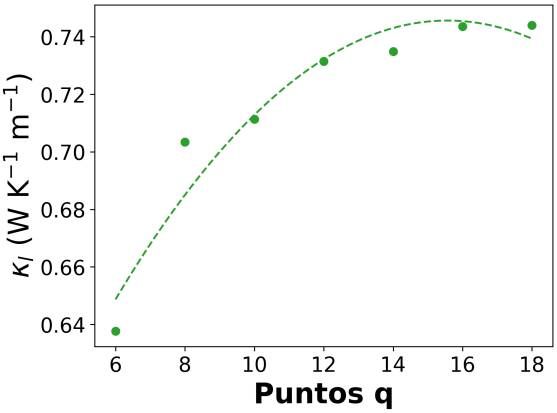

Figura 3: Convergencia de la conductividad térmica a 300 K, en función de la densidad de la malla de $N \times N \times N$ puntos **q.**

4.3 Conductividad térmica del $BaSnS_3$

Una vez validado la metodología a emplear, es posible estudiar la conductividad térmica de la perovskita $BaSnS_3$ ortorrómbica. En primer lugar se ha estudiado la conductividad térmica en función de la temperatura (Figuras 4 y 5), tanto el valor promedio como sus componentes en el tensor. Se ha explorado la conductividad desde los 200 K hasta los 1000 K en intervalos de 100 K. Se puede observar que a medida que la temperatura va en aumento, la κ_l disminuye, obteniendo un valor mínimo de $0.2\,\mathrm{W\,m^{-1}\,K^{-1}}$ a 1000 K y un valor máximo de $1.21\,\mathrm{W\,m^{-1}\,K^{-1}}$ a 200 K. Esta tendencia sigue el comportamiento de una función T^{-1} lo que demuestra que la principal fuente de fenómenos de scattering se da a través de interacciones fonón-fonón.

La Figura 4 también permite evaluar la anisotropía del material. Se aprecia que existen diferencias importantes entre las distintas componentes del tensor, siendo los valores más

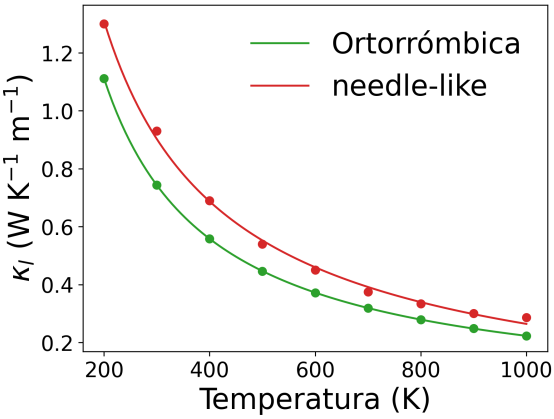

Figura 4: Conductividad térmica promedio, κ_l, en función de la temperatura.

altos la componente yy y los más bajos en la componente zz, teniendo una diferencia de casi el 67 % entre estas componentes.

Los valores de κ_l obtenidos son muy bajos, lo que pone de relieve la anarmonicidad del material. Estos ese pueden comparar con los valores obtenidos con una metodología similar a la empleada en este trabajo por Li *et al.* para la estructura de tipo "needle-like". [12] La conductividad promedio obtenida para la estrcutura ortorrómbica es ligeramente inferior a la de la estructura "needle-like". A 300 K, los valores obtenidos por Li *et al.* fueron $0.92\,\mathrm{W\,m^{-1}\,K^{-1}}$ mientras que para la perovskita ortorrómbica fue de $0.74\,\mathrm{W\,m^{-1}\,K^{-1}}$. A 900 K, estas diferencias se hacen más significativas, ya que la estructura "needle-like"presenta una κ_l de $0.32\,\mathrm{W\,m^{-1}\,K^{-1}}$ mientras que la ortorrómbica alargada $0.24\,\mathrm{W\,m^{-1}\,K^{-1}}$ (un 25 % menor).

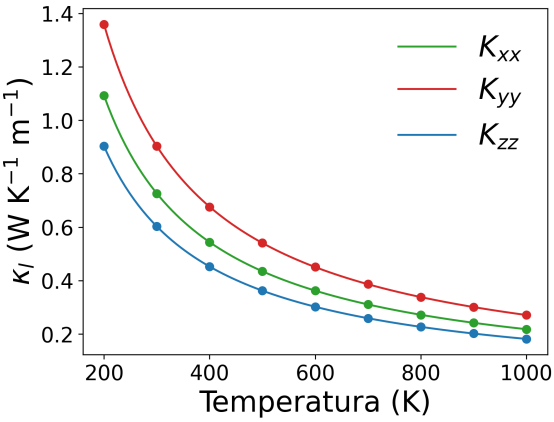

Figura 5: Componentes principales del tensor κ_l.

Los valores de κ_l tan reducidos hacen a este material interesante para su potencial

uso en dispositivos termoeléctricos, sin embargo, es necesario entender el origen de tan reducidos valores. Para ello, es importante analizar los factores que la influyen en κ_l. Tal como se describe en la sección de la metodología, la conductividad térmica depende principalmente de las frecuencias de vibración de los fonones, sus velocidades de grupo y sus tiempos de vida media.

A pesar de la baja conductividad, las velocidades de grupo no son especialmente bajas, siendo mayores para los modos acústicos, alrededor del punto Γ (ver Figura 6a). Por lo tanto, el motivo de la baja conductividad del material, debe estar relacionado con los tiempos de vida media de los fonones (ver Figura 6b). De manera general, las frecuencias de dispersión son altas, lo que justifica tiempos de vida cortos y bajas κ_l. Los únicos modos vibracionales con tiempos de vida ligeramente superiores son los módos acústicos, lo que demuestra que son los principales modos vibracionales que contribuyen a κ_l. Por lo cual, se puede concluir que los modos acústicos son los que deben contribuir en mayor medida a la conductividad térmica del $BaSnS_3$.

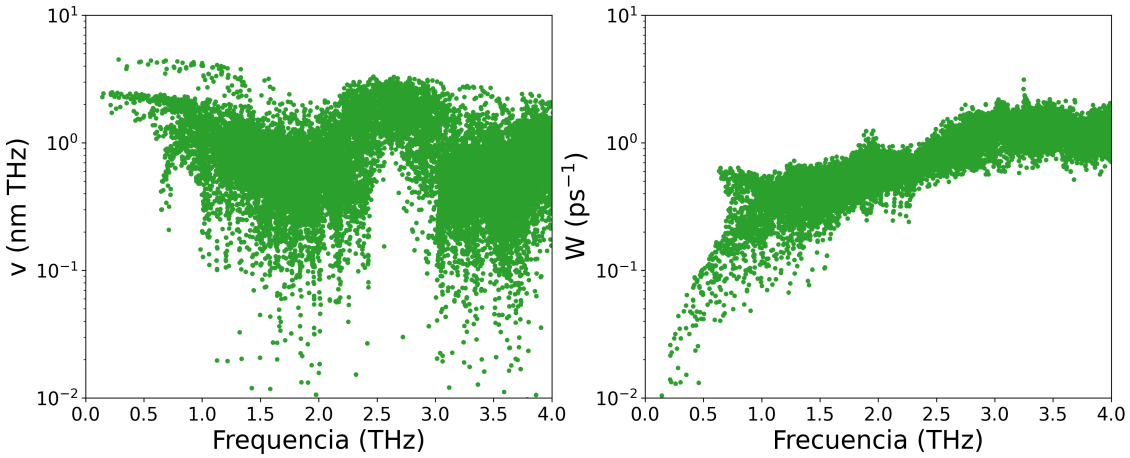

Figura 6: a) Velocidades de grupo en función de las frecuencias de vibración de cada fonón. b) Frecuencias de dispersión (inversa del tiempo de vida media) en función de las frecuencias de vibración de cada fonón.

Para poder entender mejor la anarmonicidad del material, se han represpresentado las curvas de dispersión del mismo (ver Figura 7). La ausencia de frecuencias imaginarias demuestra la estabilidad dinámica del $BaSnS_3$ en su forma ortorrómbica y su potencial síntesis como fase metaestable. Es importante destacar la gran cantidad de solapamiento entre los modos acústicos y ópticos de baja energía. Este hecho facilita la aparición de procesos de dispersión ya que es más sencillo encontrar tríos de fonones que conserven

tanto el momento como la energía, lo cual reduce κ_l.

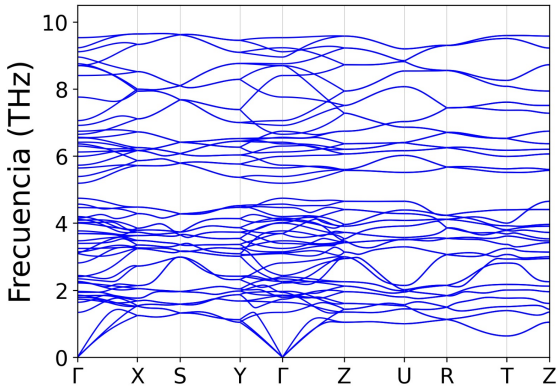

Figura 7: Curva de dispersión de fonones para el $BaSnS_3$.

Para identificar los átomos que más participan en esos modos de vibración se ha proyectado la densidad de estados de cada uno de los elementos (ver Figura 8). En esta gráfica se puede observar que el bario es el átomo que más contribuye en los modos acústicos de vibración (a menor frecuencia), mientras que el azufre es el que más aporta en los modos de vibración ópticos (a mayor frecuencia), aunque su contribución en los modos acústicos también es bastante elevada sobre una frecuencia de $3\,\mathrm{THz}$. Las vibraciones del estaño son más relevantes que las del azufre en los modos acústicos pero a partir de los $4\,\mathrm{THz}$, su contribución empieza a disminuir.

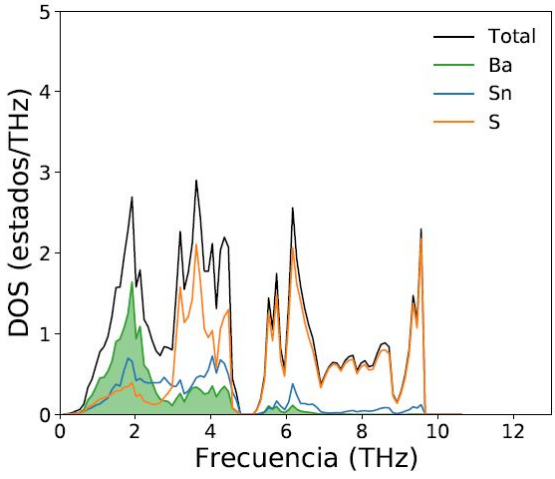

Figura 8: Densidad de estados vibracional total y proyectada en cada elemento para el $BaSnS_3$.

Hasta este punto, sólo se ha analizado κ_l, su origen, anisotropía y dependencia con la temperatura para muestras monocristalinas y libres de defectos. Sin embargo, expe-

rimentalmente, esas perovskitas, suelen ser sintetizadas como sistemas policristalinos y presentan defectos. Es por ello, que es interesante ver como varía la κ_l con el tamaño de grano en la estructura ortorrómbica.

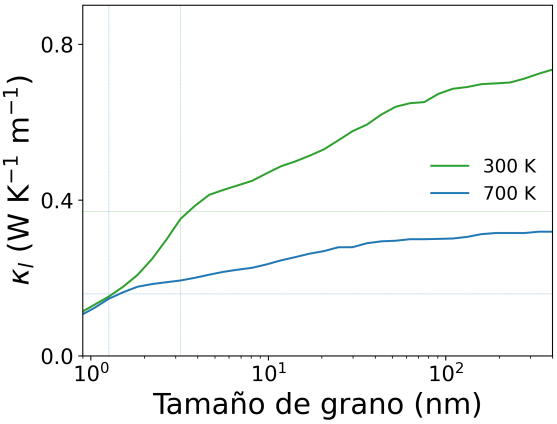

Figura 9: Dependencia de κ_l con el tamaño de grano a $300\,\mathrm{K}$ y $700\,\mathrm{K}$. Las líneas punteadas marcan el valor de tamaño de grano que reduce κ_l a la mitad del valor límite.

En la Figura 9 se puede observar la dependencia de la κ_l respecto al tamaño de grano a $300\,\mathrm{K}$ y $700\,\mathrm{K}$. Las líneas punteadas nos representan a qué tamaño de grano la conductividad térmica está en la mitad del valor límite o del sistema monocristalino. Para obtener estos valores, primero se ha calculado el valor del recorrido libre medio y a continuación, las contribuciones a κ_l de todo aquel fonón que tuviera un recorrido libre medio mayor que el tamaño del grano, han sido eliminados. El efecto del tamaño de grano es mucho más acusado a $300\,\mathrm{K}$, ya que se alcanza la mitad de valor límite de κ_l a un tamaño de grano de $3\,\mathrm{nm}$ mientras que a la temperatura de $700\,\mathrm{K}$ es necesario un tamaño de casi $1\,\mathrm{nm}$ para reducir κ_l a la mitad. Esta reducción de κ_l con el tamaño de grano es una estrategia para mejorar la eficiencia de un material termoeléctrico, mediante su nanoestructuración.

5 Conclusions

In this work, the thermal transport properties of the orthorhombic $BaSnS_3$ polymorph were explored. Although, "needle-like" structure is thermodynamically more stable than the orthorhombic polymorph, no imaginaries frequencies are found for this polymorph. This finding corroborate that orthorhombic $BaSnS_3$ polymorph is dynamically stable and can be considered a metastable phase.

Our theoretical approach was validated based on two steps. First, Machine Learning model was validated based on the validation RMSE as a function of the number of super-cells, optimization algorithm and interatomic cutoff distance. Second, κ_l was converged based on the **q** point mesh density.

Thermal conductivity was explored in a large range of termperatures (200-1000 K), finding very low values. High anisotropy was found, with the yy component of tensor being the highest values in the whole range of temperature. Orthorhombic $BaSnS_3$ polymorph presents a lower averaged κ_l if compared with the needle-like structure. Main contribution to this low κ_l comes from the large scattering rates which are a prove of the high anharmonicity of the material. This anharmoncity is due o the effective overlapping between the acoustic and low-energy optical modes. Acoustic modes are heavily projected on Ba atoms but Sn and S contributions cannot be dismissed. Nanostrcuturing seems to be a good strategy to reduce the κ_l of this material, specially at low temperarures.

The low lattice thermal conductivity of orthorhombic $BaSnS_3$ perovskite encourage i) the synthesis of this polymorph and ii) further studies focused on the electronic transport properties of this material and its ZT.

Bibliografía

[1] https://es.enerdata.net/publicaciones/informes-energeticos/tendencias-energeticas mundiales.html, "Tendencia consumo mundial energético," Consultado 03/06/2023.

[2] https://www.reportlinker.com/p06135123/Thermoelectric Market-Report-2031.html, "Thermoelectric market report 2021-2031," Consultado 03/06/2023.

[3] J. P. Heremans, S. Kalinin, and D. Morelli, "Thermoelectrics: New directions," *MRS Bulletin*, vol. 35, no. 3, pp. 188–193, 2010.

[4] G. J. Snyder and E. S. Toberer, "Efficient thermoelectrics for the 21st century," *Nature Materials*, vol. 10, no. 11, pp. 759–765, 2011.

[5] S. Datta, "Thermoelectrics," *Materials Today*, vol. 15, no. 7, pp. 20–27, 2012.

[6] J. P. Heremans, "Thermoelectrics, low-dimensionality, and nanostructured materials," *MRS Bulletin*, vol. 38, no. 10, pp. 824–830, 2013.

[7] J. C. Peltier, "Nouvelles experiences sur la caloricite des courans electriques," *Ann. Chim*, pp. 371–386, 1834.

[8] G. Snyder and A. Snyder, "Figure of merit zt of a thermoelectric device from materials properties," *Energy Environ. Sci.*, vol. 10, 09 2017.

[9] Y. Pei, H. Wang, and G. J. Snyder, "Band engineering of thermoelectric materials," *Advanced Materials*, vol. 24, no. 46, pp. 6125–6135, 2012.

[10] T. J. Zhu, Y. Liu, C. X. Fu, J. L. Xu, X. B. Zhao, and Q. Zhang, L.and Zhang, "Towards high-performance n-type thermoelectric materials by rational modification of bdppv backbones," *Energy and Environmental Science*, vol. 7, no. 12, pp. 4090–4096, 2014.

[11] P. Hohenberg and W. Kohn, "Inhomogeneous electron gas," *Phys. Rev*, vol. 136, no. 3B, pp. B864–B871, 1964.

[12] Z. Li, H. Xie, Y. Xia, and X. Tang, "Weak-bonding elements lead to high thermoelectric performance in $BaSnS_3$ and $SrSnS_3$: A first-principles study," *Journal of Materials Science*, vol. 57, no. 11, pp. 5541–5549, 2022.

[13] R. Yan, C. Shen, M. Widenmeyer, and A. Weidenkaff, "The role of interstitial Cu on thermoelectric properties of ZrNiSn half-heusler compounds," *Journal of Physics and Chemistry of Solids*, vol. 121, p. 109933, 2023.

[14] A. Majumdar, "Thermoelectrics: A materials perspective," *Annual Review of Materials Research*, vol. 43, no. 1, pp. 383–410, 2013.

[15] W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects," *Physical Review*, vol. 140, no. 4A, pp. A1133–A1138, 1965.

[16] R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules*. Oxford University Press, 1989.

[17] R. M. Martin, *Electronic Structure: Basic Theory and Practical Methods*. Cambridge University Press, 2004.

[18] J. P. Perdew and M. Ernzerhof, "Rationale for mixing exact exchange with density functional approximations," *J. Chem. Phys.*, vol. 105, no. 22, pp. 9982–9985, 1996.

[19] W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects," *Phys. Rev.*, vol. 140, pp. A1133–A1138, 1965.

[20] W. Koch and M. C. Holthausen, *A Chemist's Guide to Density Functional Theory*. Wiley-VCH, 2001.

[21] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," *Phys. Rev. Lett.*, vol. 77, pp. 3865–3868, 1996.

[22] J. J. Plata, P. Nath, J. F. Sanz, and A. Márquez, "In silico modeling of inorganic thermoelectric materials," *Comprehensive Inorganic Chemistry III*, vol. 57, no. 20, pp. 11454–11470, 2022.

[23] J. Callaway, "Model for lattice thermal conductivity at low temperatures. physical review," *Physical Review*, vol. 4, p. 1046, 1959.

[24] P. B. Allen, "Improved Callaway model for lattice thermal conductivity," *Physical Review*, vol. 14, pp. 144–146, 2013.

[25] D. A. Broido, M. Malorny, G. Birner, N. Mingo, and D. A. Stewart, "Intrinsic lattice thermal conductivity of semiconductors from first principles," *Applied Physics Letters*, vol. 91, no. 23, p. 231922, 2007.

[26] J. M. Ziman, "Electrons and phonons: The theory of transport phenomena in solids," *Clarendon Press*, pp. 1040–1050, 1960.

[27] S. Baroni, S. Gironcoli, A. D. Corso, and P. Giannozzi, "Phonons and related crystal properties from density-functional perturbation theory," *Reviews of Modern Physics*, vol. 73, no. 2, pp. 515–562, 2001.

[28] E. Eriksson, E. Fransson, and P. Erhart, "The hiphive package for the extraction of high-order force constants by machine learning," *Advanced Theory and Simulations*, pp. 180–184, 2019.

[29] T. Hastie, R. Tibshirani, and J. Friedman, "Linear methods for regression," *Springer New York*, vol. 1, pp. 43–99, 2009.

[30] M. Kuhn and K. Johnson, "Applied predictive modeling," *Springer*, vol. 1, pp. 364–489, 2018.

[31] J. Jeong and X. Chen, "Enhanced recursive feature elimination," in *2007 International Conference on Machine Learning and Applications*, (Los Alamitos, CA, USA), pp. 429–435, IEEE Computer Society, dec 2007.

[32] O. Isayev, D. Fourches, E. Muratov, C. Oses, K. Rasch, A. Tropsha, , and S. Curtarolo, "Materials cartography: Representing and mining materials space using structural and electronic fingerprints," *Chemistry of Materials*, vol. 29, no. 1, pp. 943–957, 2017.

[33] G. Kresse and J. Hafner, "Ab initio molecular dynamics for liquid metals," *Physical Review B*, vol. 47, no. 1, pp. 558–561, 1993.

[34] W. Li, J. Carrete, N. A. Katcho, and N. Mingo, "ShengBTE: a solver of the Boltzmann transport equation for phonons," *Comp. Phys. Commun.*, vol. 185, p. 1747–1758, 2014.