University of California
Santa Barbara

Accelerating the Discovery of New Energy Materials by High-Throughput Computational Search

A dissertation submitted in partial satisfaction
of the requirements for the degree

Doctor of Philosophy
in
Physics

by

Jiaoyue Yuan

Committee in charge:

Professor Bolin Liao, Committee Co-Chair
Professor Ania Jayich, Committee Co-Chair
Professor Cenke Xu

March 2024

The Dissertation of Jiaoyue Yuan is approved.

Professor Cenke Xu

Professor Ania Jayich, Committee Co-Chair

Professor Bolin Liao, Committee Co-Chair

March 2024

Accelerating the Discovery of New Energy Materials by High-Throughput Computational Search

by

Jiaoyue Yuan

Acknowledgements

First of all, I would like to thank my PhD advisor Bolin Liao for giving me the opportunity to join TEALab, for bringing me in contact with interesting minds, and for countless support and trust in my development. He motivated me with very interesting research directions and gave me complete freedom to thrive and find my own place.

I would also like to thank my devoted labmate and coauthor Yubi Chen, who shared the journey as a physicist and theorist in our group. Not only did she guide me with her expertise in first principles calculations, her enthusiasm for physics and patience as a mentor also inspired me. A big part of my knowledge about microscopic thermal and electronic properties in condensed matter was obtained in discussions with her. Our relationship certainly went beyond coworker. I am deeply grateful to have her in my PhD career.

I am much obliged to my labmates who contributed to a wonderful working environment. I would like to thank everyone in T.E.A. Lab for sharing their resilience and passion in our diverse pool of research topics. I had enjoyable and fruitful collaborations with Runqing Yang and Lokanath Patra on machine learning and magnetocaloric materials. I loved the routine dining with my labmates, and these moments will be missed. Special thanks to Runqing and Fanghao Zhang for bringing a lot of fun knowledge and experiences and being great tennis buddies. Sincere thanks to Yubi and Fanghao for supplying an infinite flow of snacks in the office. Special thanks to Wenkai Ouyang for being a great housemate in addition to a labmate.

I would like to express my deepest gratitude and appreciation to my dearest room-

mate and closest friend Swaroopa Lahiri and the best cat in the observable universe Blackberry. Their persistent support and trust have been with me throughout my entire PhD life and will continue to stay around as I embark on new adventures. Swaroopa is not only my best friend, but also an intellectual equal from whom I learned so much about social science and humanities. Our endless discussions on global inequality and marginalized communities will always be a reminder for me to strive for a better world for all. Blackberry, a true emotional support animal, fulfilled her important role to keep me mentally healthy. Her cute little face and fluffy white paws comforted my soul during all the hardships.

I extend my heartfelt thanks to my close friends who shared life with me and shaped me into who I am. They showed me how to be loved, how to love others, and how to love one's job. I am inspired by friends who are able to supply energy sustainably across professional commitments and personal interests. Special thanks to my best friends who have been here for me even when they were on the other side of the globe. I am also thanking my partner Georg Winkler for having quality fun time with me during the most stressful and busiest times and not commenting on my research. They made this untamable spirit last longer.

Lastly, and most importantly, I am profoundly grateful to have my parents for their almost unconditional love, respect, full support and 100% trust throughout my life.

Thank you all for being a part of my PhD journey.

Curriculum Vitæ

Jiaoyue Yuan

Education

2024	Ph.D. in Physics (Expected), University of California, Santa Barbara.
2020	M.A. in Physics, University of California, Santa Barbara.
2017	B.S. in Astrophysics, Rutgers University, New Brunswick.

Publications

Jiaoyue Yuan, Runqing Yang, Lokanath Patra, Bolin Liao, *"Enhancing Magnetocaloric Material Discovery: A Machine Learning Approach Using an Autogenerated Database by Large Language Models"*, arXiv:2403.02553, (submitted to Journal of Applied Physics, under review, 2024).

Jiaoyue Yuan, Yubi Chen, Bolin Liao, *"Lattice Dynamics and Thermal Transport in Semiconductors with Anti-Bonding Valence Bands"*, Journal of American Chemical Society, 145 (33), 18506-18515 (2023).

Fields of Study

Physics (Advisor: Bolin Liao)

Abstract

Accelerating the Discovery of New Energy Materials by High-Throughput Computational Search

by

Jiaoyue Yuan

The quest for sustainable, efficient energy systems is more pressing than ever as we step into the 21st century. Finding the suitable materials for a specific technology is a very challenging task. The required properties can be unique to particular chemical compositions and structures. Furthermore, achieving high performance often demands the precise alignment of multiple interacting properties, which can have competing effects. Additional economic constraints, such as the cost of raw materials and the manufacturing process, further add to the challenges. Traditionally, this journey has been arduous, relying on slow and laborious experimental techniques. One pivotal approach that has revolutionized the pace and efficiency of this discovery process is High-throughput screening (HTS). HTS is a powerful and systematic methodology that enables the rapid evaluation of a large number of materials in a relatively short period by leveraging large databases, accurate and efficient computational techniques including *ab initio* calculations and machine learning. When large libraries of materials are screened based on specific properties or descriptors, only a limited amount of the most promising ones pass to the final stage for further validation through expensive calculations and experiments. This thesis aims to develop HTS pipelines on finding promising candidates for two energy applications: thermoelectric materials that convert heat into electricity, allowing for waste heat recovery, and magnetocaloric materials that serve an environmental friendly cooling technology. The integration of predictive models and screening pipelines signif-

icantly reduces the time and cost associated with material discovery, transforming the traditional methods of discovering and developing novel materials.

In the pursuit of innovative solutions for sustainable energy, thermoelectric materials hold undeniable potential to transform energy generation. They possess the unique ability to convert thermal energy into electricity, offering promising avenues for harnessing waste heat and enhancing energy efficiency. Achieving high thermoelectric performance requires efficient manipulation of thermal conductivity and a fundamental understanding of the microscopic mechanisms of phonon transport in crystalline solids. One of the major challenges in thermal transport is achieving ultralow lattice thermal conductivity. We use the anti-bonding character of the highest-occupied valence band as an efficient descriptor for discovering new materials with an ultralow thermal conductivity. We first examine the relationship between anti-bonding valence bands and low lattice thermal conductivity in model systems PbTe and $CsPbBr_3$. Then, we perform a high-throughput search in the Materials Project database and identify over 600 experimentally stable binary semiconductors and report the anti-bonding strength in their valence bands. From our candidate list, we conduct a comprehensive analysis of the chemical bonds and the thermal transport in the XS family, where X=K, Rb, and Cs are alkaline metals. These materials all exhibit ultralow thermal conductivities less than 1 W/(m K) at room temperature despite simple structures. We attribute the ultralow thermal conductivity to the weakened bonds and increased phonon anharmonicity due to their anti-bonding valence bands. Our results provide chemical intuitions to understand lattice dynamics in crystals and open up a convenient venue towards searching for materials with an intrinsically low lattice thermal conductivity.

In another search for new energy materials, magnetocaloric materials are requested as promising candidates for energy-efficient and eco-friendly cooling technologies. These materials exhibit the magnetocaloric effect, where their temperature changes in response

to the addition or removal of an external magnetic field. This unique property opens doors to environmentally friendly refrigeration systems since it eliminates the need for refrigerants with high global warming potential. Magnetic cooling based on the magnetocaloric effect is a promising solid-state refrigeration technology for a wide range of applications in different temperature ranges. Previous studies have mostly focused on near room temperature (300 K) and cryogenic temperature (<10 K) ranges, while important applications such as hydrogen liquefaction call for efficient magnetic refrigerants for the intermediate temperature 10 K to 100 K. For efficient use in this range, new magnetocaloric materials with matching Curie temperatures need to be discovered, while conventional experimental approaches are typically time-consuming and expensive. In this work, we report a computational material discovery pipeline based on a materials database containing more than 6000 entries auto-generated by extracting reported material properties from literature using a large language model. We then use this database to train a machine learning model that can efficiently predict magnetocaloric properties of materials based on their chemical composition. We further verify the magnetocaloric properties of predicted compounds using *ab initio* atomistic spin dynamics simulations to close the loop for computational material discovery. Using this approach, we identify 11 new promising magnetocaloric materials for the target temperature range. The joined force of AI advances and high-throughput search pipelines enables more efficient prediction and iteration. Together they are revolutionizing the conventional ways to discover new materials.

Contents

Chapter 1

Introduction

1.1 Navigating the Future of Energy

As we step into the 21st century, it's widely recognized that the future energy landscape faces numerous challenges. The quest for sustainable, efficient, and accessible energy sources is more pressing than ever, given the looming threats of climate change, dwindling fossil fuel reserves, and the escalating demand for energy driven by global population growth and industrialization. The world's current energy infrastructure, heavily reliant on fossil fuels, is unsustainable. The burning of coal, oil, and natural gas is a major contributor to greenhouse gas emissions, leading to climate change and environmental degradation. Moreover, they may no longer be viable as primary energy sources due to the finite nature of these resources. This scenario requires a pivotal shift towards renewable energy sources such as solar, wind, hydro, and bioenergy. However, the transition is fraught with technological, economic, and logistical challenges, many of which are deeply intertwined with the materials used to harness, store, and transmit energy.

At the heart of solving these multifaceted energy challenges lies the field of material science, a domain that holds the keys to unlocking revolutionary energy technologies.

Whether it's improving the efficiency of solar panels, extending the life of batteries, or developing new materials for sustainable refrigeration systems, the advances in material science directly influence the viability and performance of renewable energy solutions. The ability to engineer materials at the atomic or molecular level opens up unprecedented possibilities for energy generation, storage, and conservation.

Solar energy is one of the most promising renewable sources. The efficiency of solar cells largely depends on the materials used to convert sunlight into electricity. For decades, silicon has been the backbone of photovoltaic (PV) technology, primarily due to its abundance, non-toxicity, and well-understood properties. Silicon solar cells have seen significant improvements in efficiency, with commercial products now reaching about 20-25% [2]. However, the intrinsic limitations of silicon, such as its indirect bandgap, requires thick layers to absorb sufficient sunlight, leading to higher material and manufacturing costs. Moreover, the efficiency of silicon solar cells is nearing its theoretical maximum [3], pushing researchers to explore alternative with higher potential. Perovskites have emerged as a highly promising class of materials for next-generation solar cells. Characterized by their crystal structure, perovskite materials offer a direct bandgap, enabling them to absorb light more efficiently than silicon and potentially achieve higher conversion efficiencies with significantly thinner layers. Additionally, perovskite solar cells can be manufactured using low-cost solution-based processes. Despite their advantages, perovskite cells face challenges related to durability, toxicity (due to lead content), and long-term environmental stability [4], which are active areas of research. Other emerging materials include organic photovoltaics (OPVs) and tandem solar cells. OPVs leverage carbon-based compounds to create flexible, lightweight solar cells that offer potential for lower production costs and adaptable applications, although they currently face challenges in efficiency and longevity [5]. Tandem solar cells, combining layers of different semiconducting materials optimized to absorb various segments

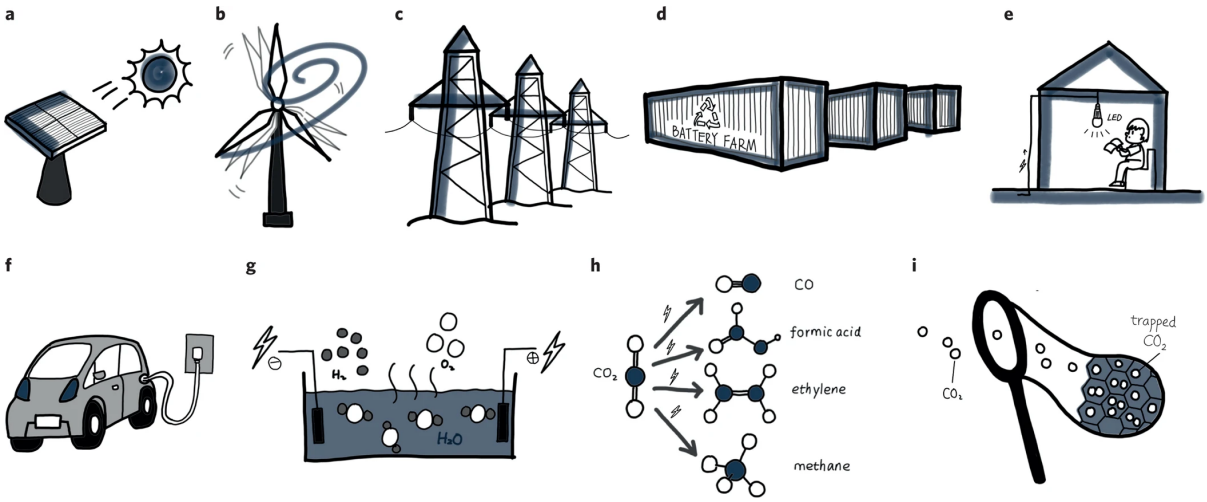

Figure 1.1: **Development in energy systems from clean renewable energy sources.** (a) Solar panels. (b) Wind turbines. (c) Electricity distribution cables. (d) Battery as energy storage system. (e) Energy management in buildings. (f) Electric vehicles. (g) Catalysts and photoelectrochemical devices. (h) CO_2 conversion to fuels and chemicals. (i) Carbon-capture. Illustration courtesy of Paloma Liu. [1] Reused with the permission from Springer Nature.

of the solar spectrum, aim to exceed the efficiency limits of single-junction cells, with silicon-perovskite tandems showing particularly promising results [6]. With ongoing research in material science, the challenges must be addressed before these new materials can be widely adopted.

Battery is another energy technology that urges the advances in material science. The demand for high-capacity, long-life batteries is soaring, driven by the rise of electric vehicles (EVs) and the need for energy storage solutions for renewable energy systems. Current lithium-ion (Li-ion) batteries are limited by the availability of lithium, cost, and environmental concerns. Material scientists are exploring alternatives like sodium-ion batteries, solid-state batteries, and lithium-sulfur batteries, which promise higher energy densities, safety, and lower environmental impact. However, the reliance on lithium and cobalt, materials that are limited in supply and associated with environmental and ethical concerns, along with safety issues related to thermal runaway, highlight the need for alternative materials and technologies [7]. Material science research is actively exploring alternatives to traditional anode and cathode materials to enhance battery performance and safety. Silicon anodes, for example, can store up to ten times more lithium than graphite anodes used in conventional Li-ion batteries, offering a significant boost in energy density [8]. Meanwhile, cathode advancements include the development of high-nickel chemistries that reduce cobalt content, enhancing energy density while mitigating ethical and supply chain concerns [9]. Sodium-ion (Na-ion) batteries emerge as a promising alternative, leveraging the abundant and widely available sodium element [10]. A transformative leap would be solid-state batteries, replacing the liquid electrolytes in conventional batteries with solid materials. Overcoming challenges related to interface stability and manufacturing scalability remains a key focus in the field of research.

Thermoelectric materials possesses an unique feature for power generation, capitalizing on the direct conversion of temperature differences to electrical voltage and vice

versa. This category of materials contribute to sustainable energy solutions for their potential in capturing and converting waste heat from industrial processes, automotive exhaust, and even ambient environmental heat into usable electricity. The field of thermoelectrics directly benefits from the exploration and development of materials that can efficiently conduct electricity while minimally conducting heat. This delicate balance is crucial for enhancing the thermoelectric figure of merit (ZT), a dimensionless measure of a material's efficiency in converting heat into electrical energy. $ZT = \frac{\sigma S^2 T}{\kappa}$, where T is the temperature, S is the Seebeck coefficient, which quantifies a material's ability to generate a voltage in response to a temperature gradient, σ is electrical conductivity, and κ is thermal conductivity, which has electron contribution κ_e and lattice contribution κ_L. The development of thermoelectrics materials is enabled by discovering and engineering materials with high Seebeck coefficients and electrical conductivity, coupled with low thermal conductivity [11, 12, 13]. This often involves complex doping strategies, nanostructuring, and the exploration of inherently low thermal conductivity materials [14]. Doping, the addition of impurities to a pure semiconductor, is a common strategy used to adjust the carrier concentration, thereby optimizing the electrical conductivity and Seebeck coefficient. Nanostructuring is a powerful technique to reduce thermal conductivity without adversely affecting electrical properties. By introducing boundaries and interfaces at the nanoscale, phonons (heat carriers) scatter more effectively, reducing thermal conductivity. This can involve the creation of nano-sized grains, interfaces, or the inclusion of nanoparticles within the bulk material [15, 16, 17, 18]. In addition, discovering new thermoelectric materials with inherently favorable properties is another direction of research. This includes materials with complex crystal structures that naturally scatter phonons, compounds with low-dimensional structures, and materials exhibiting strong electron-phonon coupling [19, 20]. High-throughput computational screening and machine learning are increasingly employed to predict and identify

promising thermoelectric materials, accelerating the discovery process [21, 22, 23, 24].

Magnetic refrigeration is an innovative and environmentally friendly cooling technology that relies on the magnetocaloric effect (MCE). The MCE occurs when a magnetic material changes temperature upon the application or removal of a magnetic field. When the material is placed in a magnetic field, its magnetic moments align with the field. This alignment is a more ordered state and results in a decrease in the material's magnetic entropy. The change in entropy must be compensated by a change in thermal energy to maintain thermodynamic equilibrium. As a result, when the entropy decreases (as the material orders magnetically), the material absorbs heat from its surroundings, causing it to cool down [25]. Conversely, when the magnetic field is removed, the magnetic moments become disordered, increasing the material's entropy. To balance this, the material releases heat to its surroundings, leading to an increase in temperature. In a magnetic refrigeration system, this effect is used in a thermodynamic cooling cycle. This offers a potential alternative to traditional vapor-compression refrigeration systems, which require refrigerants that exacerbate global warming. The advancement for efficient magnetic refrigeration begins with identifying materials that exhibit a significant MCE. The efficiency of MCE is often measured by the amount of temperature change the material undergoes with the application or removal of a magnetic field and the total heat transfer capacity of the system, which are determined by two key factors: Curie temperature (T_C) and Magnetic Entropy Change (ΔS_M). T_C is a critical point for magnetic materials where they undergo a transition from a magnetically ordered state to a magnetically disordered state. ΔS_M directly relates to the amount of heat absorbed or released by the material when magnetized. So larger ΔS_M means higher cooling capacity. T_C determines the optimal working temperature for an MCE material, since ΔS_M are typically maximized near T_C due to a sharp change in magnetization. Various classes of materials have been explored to find candidates with large ΔS_M at different temperature ranges.

For example, rare earth-based materials demonstrate excellent magnetocaloric properties at low operating temperature (T_C around 10 K to 300 K), while transition metal-based alloys show higher T_C and are more cost-effective compared to rare earth-based materials [26]. In addition to finding the suitable refrigerants, more challenges including cost for generating strong magnetic field, system integration and scaling, efficient heat transfer systems need to be addressed for widespread adoption of magnetic refrigeration [27].

The journey to overcome these material challenges is complex and multidisciplinary, involving physicists, chemists, engineers, and environmental scientists. The discovery of new materials and the optimization of existing ones are taking actions to address the global energy challenge. As this chapter has introduced, innovations in solar cells, battery storage systems, thermoelectric materials, and magnetic refrigeration and beyond are pivotal for the advancement of clean energy solutions. It is within this context that the focus of this thesis narrows, concentrating specifically on the search for new promising candidates suited for thermoelectric applications (Sec. 1.3) and magnetic refrigeration (Sec. 1.4).

1.2 High-throughput Search in Material Science

High-throughput Search (HTS) in material science represents a transformative approach to the discovery and optimization of new materials, leveraging the power of automation and computational techniques to rapidly screen vast libraries of compounds. This methodology marks a paradigm shift from traditional, slow, and often serendipitous material discovery processes towards a more systematic, efficient, and predictive framework. The foundation of HTS is built upon the integration of computational methods, such as density functional theory (DFT), machine learning algorithms, and availability of large databases. Computational methods provide the ability to predict material

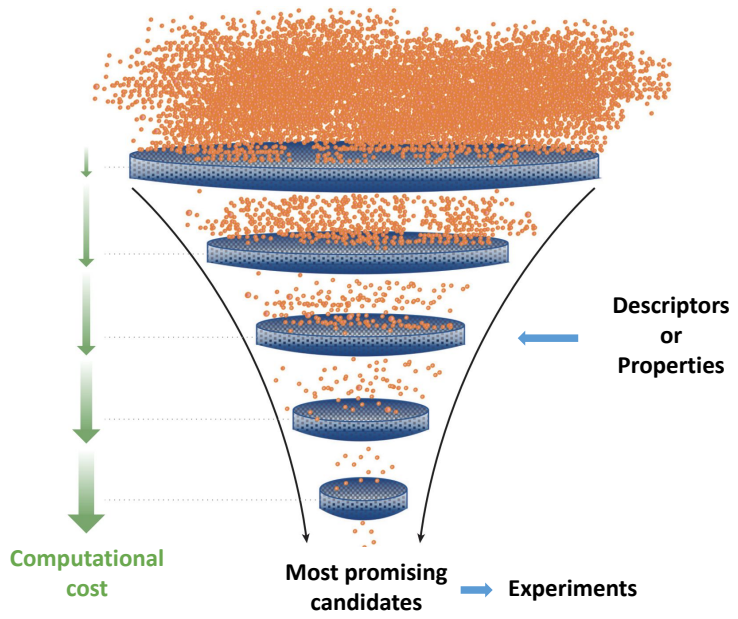

Figure 1.2: **Tiered high-throughput search in computational material science.** This process involves a funnel-like approach to evaluate properties or descriptors, starting with simpler, minimal computations and progressively moving to more complex and computationally demanding ones. Adapted with permission from Annual Review of Materials Research.

properties and identify promising candidates by simulating their electronic, mechanical, and thermal behaviors at the atomic level. Machine learning algorithms further accelerate the searches by learning from existing data, improving prediction accuracy, and most importantly reducing the computational cost associated with simulations. Large databases are make available by open access resources and data mining techniques that can sift through extensive unstructured data. Together, the pipeline can rapidly narrow down the vast space of possible materials to a manageable number of candidates for experimental synthesis and characterization, thereby streamlining the discovery process.

Recent studies based on high-throughput search methodologies have significantly advanced our understanding and discovery of materials with promising applications in various domains. One of the very first studies that utilized HTS focused on thermoelectrics

material and thermal conductivity management. This study combined the inorganic crystal structure database and an automated band structure analysis using augmented plane wave method. After performing *ab initio* calculations for 570 compounds screened from 1,630 Sb containing compounds in the Inorganic Crystal Structure Database, LiZnSb was identified as a potential thermoelectric material with excellent anisotropic lattice thermal properties [22]. In another study, the search for thermoelectric materials with inherently low lattice thermal conductivity led to the discovery of Cu3ErTe3 through screening of over 27,782 entries from the Materials Informatics Platform, demonstrating the material's low sound speed and promising thermoelectric performance at high temperatures [23]. Additionally, an innovative approach employing machine learning algorithms and automatic *ab initio* calculations has been applied to estimate the thermal conductivity of 450 half-Heusler compounds out of 79,000 entries in the AFLOWLIB.org database, revealing a wide variance in thermal conductivities and identifying candidates with exceptionally low values for experimental validation [24]. High-throughput *ab initio* computing also contribute to other areas of research in material discovery. For example, in the quest for superior lithium intercalation cathodes, one study has spotlighted new mixed polyanions compounds by HTS, suggesting them as candidates for high-capacity and specific energy cathode materials for lithium-ion batteries [28]. In another study looking for magnetic refrigerants, HoB_2 was found to be one of the best candidates so far through data mining, machine learning, *ab initio* calculations and experimental verification [29]. These are a few examples of many HTS studies that combine the power of large databases, *ab initio* calculations, and machine learning. Especially with the development and widespread adoption of machine learning techniques since 2012, the computational cost in HTS pipelines can be further reduced by incorporating state-of-the-art machine learning methods from statistical models to artificial neural networks. Collectively they highlight the remarkable potential of HTS techniques in material science, accelerating the

discovery and optimization of materials for critical applications in energy systems. While computational approaches are at the heart of high-throughput searches, their integration with automated experimental techniques is crucial for validating predictions and refining models. Recently, an autonomous system called A-Lab, which integrates robotics and artificial intelligence (AI), has unveiled its first set of material discoveries with potential applications in batteries and solar cells. [30]. Concurrently, another AI system developed by Google has identified hundreds of thousands of stable material candidates for future exploration by the A-Lab [31]. The integration of AI with high-throughput screening in robotic systems marks a thrilling advancement in the field of materials science and beyond. This fusion accelerates the pace of discovery not only in the predictive computational methods, but also allows automation in experimental synthesis and analysis. With AI-driven insights guiding the experimentation process, robotic systems can execute complex tasks with remarkable efficiency and minimal human intervention.

Another recent technology with the potential to revolutionize a wide range of scientific domains is Large Lange Models (LLMs). With their remarkable capabilities in processing and generating human-like text, LLMs are capable of providing information across a wide range of topics. Especially with proper prompt engineering (craft effective input text, i.e., prompts) to guide LLMs (and generally AI systems), they are capable of generating desired outputs tailored to specific tasks. The integration of LLMs, such as GPT-4, into the framework of high-throughput search in material science opens up opportunities for extracting information from vast amounts of unstructured data for database creation or property prediction. Their potentials extend beyond materials science to various scientific disciplines [32]. These AI advancements introduces a new era of scientific exploration and technological progress.

1.3 Thermoelectric Materials

Thermoelectric materials offer a fascinating glimpse into the future of energy, holding the promise to convert waste heat into electricity. The fundamental physics of thermoelectric materials revolves around the Seebeck effect, which enables the conversion of a thermal gradient into an electric voltage. This phenomenon arises due to the movement of charge carriers in response to temperature differences within the material. The intrinsic qualities of thermoelectric materials, including electrical conductivity (σ), thermal conductivity (κ), and the Seebeck coefficient (S), collectively govern their thermoelectric efficiency $ZT = \frac{\sigma S^2 T}{\kappa}$, known as the thermoelectric figure of merit, where T is the temperature. To maximize this value, one major challenge is to minimize the lattice thermal conductivity κ_L while maintaining the electrical conductivity σ. Previous studies have successful developed techniques to manipulate phonon transport by introducing defects and/or interfaces [33, 15, 16, 17, 18]. However, these extrinsic approaches can also interfere with electron transport [14], thus compromising the efficiency. For this reason, crystalline materials with an intrinsic ultralow thermal conductivity are particularly desirable.

The conversion between thermal and electrical energy makes thermoelectric materials valuable for both power generation and temperature control. The thermoelectric energy conversion is illustrated in Fig. 1.3. The Seebeck effect allows them to transform a substantial amounts of waste heat, generated by various industrial processes such as automotive, manufacturing, and power plants, into usable electrical energy. They are also suitable for space missions to generate electricity from heat sources in limited space and maintenance capabilities. The reverse of Seebeck effect, known as the Peltier effect, can be utilized for environmentally friendly cooling applications. The Peltier effect involves the absorption or release of heat when an electric current passes through a junction

of two dissimilar conductors, enabling active heating or cooling. This type of solid-state refrigeration offers a promising alternative to traditional vapor compression systems for key advantages in the compactness, environmental considerations, and precision in temperature control [34]. Thermoelectric modules are typically compact and lightweight, making them suitable for applications where space and weight are critical considerations. Additionally, they require no moving parts, whereas traditional refrigeration systems use compressors and fluids and thus can be prone to mechanical failures. Some thermoelectric materials are non-toxic, and the absence of refrigerants with high global warming or ozone depletion potential is an advantage for long term environmental goal. Lastly, the temperature control in thermoelectric refrigeration is highly responsive, allowing for precise and rapid adjustments to the cooling process. This feature is valuable in applications where maintaining a specific temperature is critical, such as in laboratories or medical storage.

The field of thermoelectric materials is multifaceted, encompassing a range of compounds with unique properties tailored for specific applications in targeted temperature ranges. Bismuth Telluride (Bi_2Te_3) is one of the most well-known and widely used thermoelectric materials for room temperature applications. It has low thermal conductivity of around 1.5 W/mK at room temperature, which is attributed to strong phonon anharmonicity [35]. The optimal ZT of Bi_2Te_3 and its alloys can exceed 1 [21, 17, 36]. Lead Telluride (PbTe) is another highly efficient thermoelectric material, especially at medium to high temperatures (500 K to 900 K). Its efficiency is attributed to its narrow bandgap, high carrier mobility, and intrinsic low thermal conductivity [37]. Its thermal conductivity is relatively low, about 2.2 W/mK at room temperature, and the ZT value can reach 2 by alloying. PbTe and its alloys are used in power generation applications, particularly for recovering waste heat from industrial processes or automotive exhausts. Another material with even higher ZT value of 2.6 in a similar temperature range (500-

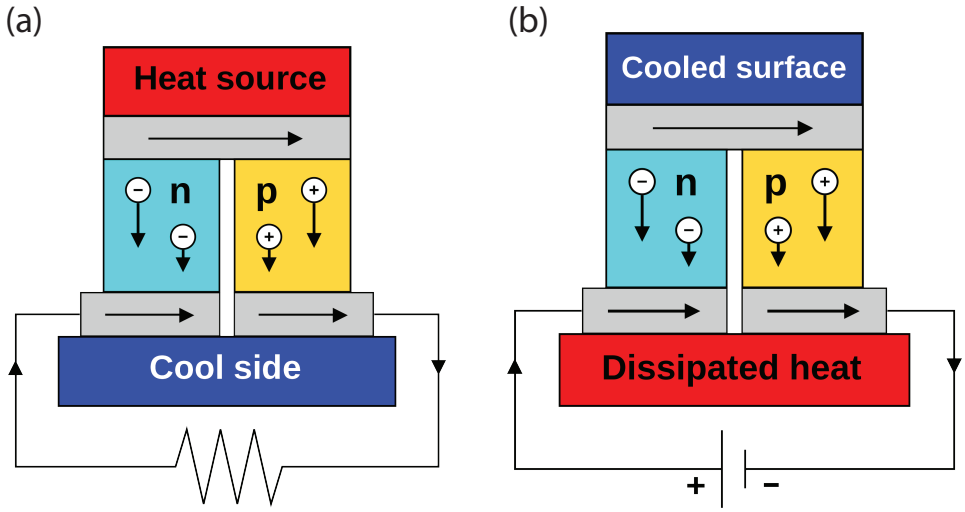

Figure 1.3: **Illustration of the principles of thermoelectric material applications.** (a) **Seebeck Effect:** The illustration shows a thermoelectric material connected to a load resister, with one end heated (red) and the other end cooled (blue), demonstrating the electrical voltage due to the temperature gradient. This effect is the basis for thermoelectric generators (TEGs) that convert waste heat into electrical energy. (b) **Peltier Effect:** This panel highlights the working principles of thermoelectric coolers (TECs), with electrical current flowing through a thermoelectric material, causing one junction to absorb heat (blue) and the other to release heat (red), thus creating a cooling effect. This principle is utilized in cooling applications for electronic devices Reproduced with the permission of Free Software Foundation, Inc. under GNU Free Documentation License.

900 K) is Tin Selenide (SnSe). This high ZT value can be credited to an ultralow thermal conductivity of 0.2-0.7 W/mK [38]. Research has focused on enhancing the ZT value of these group IV materials through doping, introducing nanostructures, and creating phonon scattering centers to further reduce thermal conductivity [39, 40, 41, 42]. Skutterudites, with a general formula of MX_3, where M is a metal such as Co, Rh, or Ir, and X is a pnictogen such as P, As, or Sb, are notable for their "phonon-glass electron-crystal" (PGEC) behavior [43]. This characteristic allows them to have low thermal conductivity due to strong phonon anharmonicity while maintaining good electron mobility. Their thermal conductivity can be exceptionally low, often below 1 W/mK, due to the phonon scattering by heavy atom fillers or voids in their structure. This low thermal conductivity is a significant factor in their potential for high thermoelectric efficiency [21, 13]. Silicon Germanium (SiGe) alloys and half-Heusler compounds (XYZ, where X and Y are metals and Z is a semiconductor) are critical in high-temperature thermoelectric applications, such as space power generation [16]. A ZT of around 1.0 near 1000 K has been achieved for both n-type and p-type half-Heusler compounds compounds [44]. Although the bulk ZT value for SiGe is lower than that of some other materials, it can be boosted to 0.5 or even up to 2 by alloying in nanowire structures [45]. Both materials combine good thermoelectric properties with high mechanical strength and thermal stability, making them suitable for use in harsh environments. Researchers in thermoelectric materials focus on reducing thermal conductivity while maintaining and/or enhancing electrical properties. Techniques such as nanostructuring, alloying, and the development of composite materials are explored to achieve lower thermal conductivities and higher figure of merit ZT values [21].

The ongoing research and development in this field are not only focused on improving the efficiency of existing materials but also on discovering new materials with higher performance and environmental compatibility. Emerging technologies including topological

insulators and quantum dot-based materials are gaining attention for their potential to push the boundaries of thermoelectric efficiency [46, 47, 48]. Organic and hybrid materials are explored for cost effectiveness and sustainability compared to traditional inorganic thermoelectric materials [49]. In this quest, HTS through the incorporation of machine learning and computational methods holds promise in identifying optimal compositions and structures for thermoelectric materials. In Chapter 2, I will introduce our HTS work that uses anti-bonding valence band as an effective descriptor to identify materials with intrinsic ultralow thermal conductivity.

1.4 Magnetocaloric Materials

In the search for sustainable and efficient cooling technologies, magnetocaloric materials emerges as a revolutionary force. These materials enable refrigeration for a wide range of applications in different temperature ranges through the magnetocaloric effect (MCE), where magnetic material changes temperature when exposed to a changing magnetic field. When an external magnetic field is applied, the material's magnetic moments (atomic magnetic orientations) tend to align with the field. This alignment reduces the material's magnetic entropy, and as a result, it releases thermal energy, causing the material to heat up [25]. Conversely, when the magnetic field is decreased, the magnetic moments in the material become disordered again. This increases the magnetic entropy, and thermal energy needs to decrease under adiabatic conditions, causing it to cool down. When the magnetization and demagnetization processes alternate, a thermodynamic cooling cycle is formed. Unlike other cooling methods that rely on harmful refrigerants or moving parts, magnetic refrigeration stands out for its solid-state and cryogen-free nature, large Coefficient Of Performance (COP), reliability and compactness [27]. These advantages make magnetic refrigeration attractive for space applications, scientific instruments, and

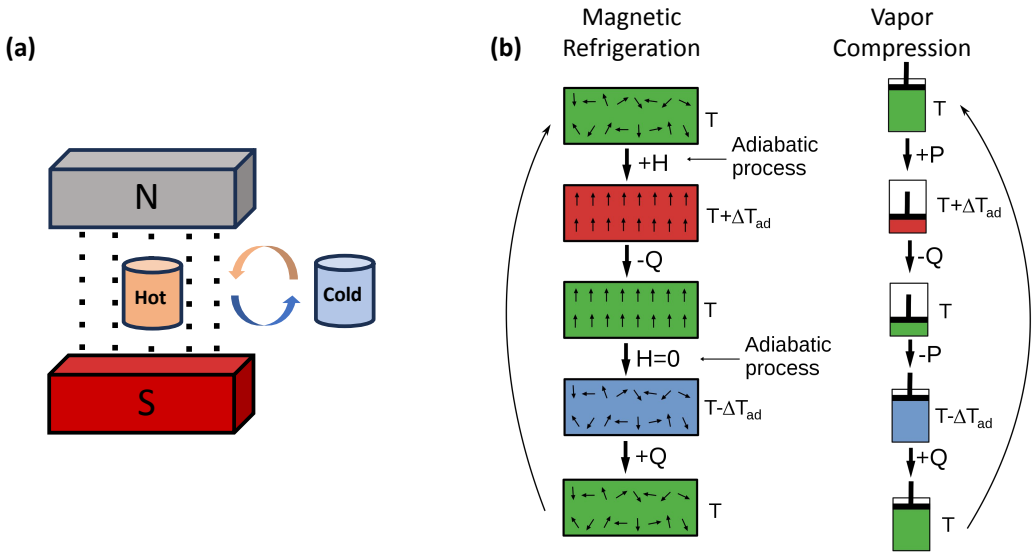

Figure 1.4: **Working principles of magnetic refrigeration.** (a) Sketch of a magnetocaloric material alternately absorbs and releases heat as it is exposed to and removed from an external magnetic field, leading to a temperature change. (b) Analogy between magnetic refrigeration and conventional vapor compression cooling. H = external magnetic field; Q = heat; P = pressure; ΔT_{ad} = adiabatic temperature change during operation. In a working cycle, this material first absorbs heat from the environment, cooling the surrounding area when magnetized, and then expels heat upon demagnetization, which is carried away by a heat exchange fluid, completing the cooling cycle. Panel (b) is licensed under the Creative Commons Attribution-Share Alike 4.0 International license.

as an eco-friendly alternative to conventional cooling such as vapor compression refrigeration. A comparison of the work cycles of magnetic refrigeration and vapor compression refrigeration is illustrated in Fig. 1.4. On the other hand, the primary disadvantages of magnetic refrigeration include high initial cost due to scare rare earth elements and magnets to maintain strong magnetic fields, heat transfer optimization for multi-stage systems to achieve a wide temperature range [50, 27].

As shown in the S-T diagram in Fig. 1.5, the effectiveness of a magnetocaloric material is determined by two key factors: Magnetic Entropy Change (ΔS_M) and Adiabatic

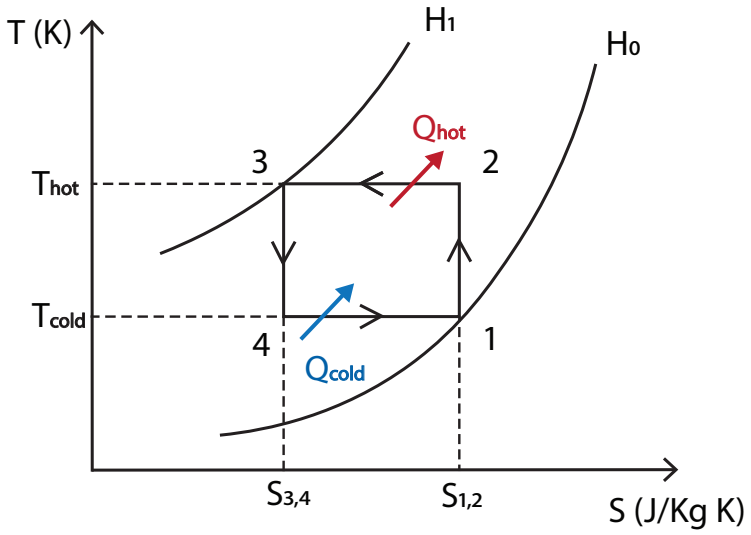

Figure 1.5: **S-T diagram of a thermodynamic cycle of magnetic refrigeration.** $1 \rightarrow 2$ Adiabatic magnetization: An external magnetic field is applied, causing the magnetic moments to align. Decreased magnetic entropy enforces an increase in lattice entropy, resulting in an adiabatic temperature rise. $2 \rightarrow 3$ Isothermal Magnetization: Magnetic field is further enhanced until the material reaches total magnetization. Heat exchange is allowed to dump waste heat under constant temperature. $3 \rightarrow 4$ Adiabatic Demagnetization: The magnetic field is gradually removed, causing the magnetic moments to be more disordered. The material cools down adiabatically. $4 \rightarrow 1$ Isothermal Demagnetization: The work cycle is completed until the material is fully demagnetized. The material absorbs heat from the refrigerated environment. The cycle starts again when thermal equilibrium is achieved.

Temperature Change (ΔT_{ad}), which can be calculated by

$$\Delta S_M = \mu_0 \int_{H_i}^{H_f} \left(\frac{\partial M}{\partial T} \right)_H dH, \tag{1.1}$$

$$\Delta T_{ad} = \mu_0 \int_{H_i}^{H_f} \frac{T}{C} \left(\frac{\partial M}{\partial T} \right)_H dH. \tag{1.2}$$

Temperature- and field-dependent magnetization and specific heat are needed in the calculation. ΔS_M and ΔT_{ad} quantify the amount of heat absorbed or released during the application or removal of a magnetic field. High values of both parameters suggest a material that can achieve significant cooling with less energy input, making the refrigeration process more efficient. A sharp change of magnetization with respect to temperature is expected for high values. For this reason, ΔS_M and ΔT_{ad} are generally maximized near Curie temperature (T_C). T_C is the temperature at which a ferromagnetic or ferrimagnetic material undergoes a phase transition, typically from a magnetized state to a non-magnetized state. Thus, T_C determines the ideal working temperature for an MCE material.

In this context, large magnetic moments are favorable characteristics for good MCE candidates, such as f-metals (Gd, Tb, Dy, Ho, Er, Tm, Eu, Sm) or d-metals (Mn, Fe, Co, Ni, Cr) with a partially filled f or d band [51]. The state-of-the-art MCE candidates include manganese-based compounds and rare-earth-based alloys. Especially the later exhibits strong magnetic interactions and can be tuned for specific temperature ranges. For example, heavy rare-earth-based MCE materials, such as HoB_2 [29, 52], $ErCo_2$ [53], $DyAl_2$ [54, 55], and $ErAl_2$ [54] show significant ΔS_M, and they are actively researched for applications like hydrogen liquefaction because they have T_C within the required range of around 10 K to 100 K [56, 57, 58]. Room temperature candidates such as ferromag-

netic perovskite manganites $R_{1-x}M_xMnO_3$, where R=La, Nd, Pr and M=Ca, Sr, Ba, etc are studied systematically for their low material processing cost compared to other magnetocaloric compounds [59].

Previously, the identification of MCE materials primarily involved an experimental approach encompassing synthesis and characterization. This process tends to be costly, time-intensive, and frequently depends on a trial-and-error methodology. To push forward, utilizing HTS and computational modeling can accelerate the identification of novel materials with superior MCE properties. Recent studies that incorporated density functional theory (DFT) and machine learning-based methods were able to predict and optimize MCE performance from physical properties while reducing material cost [60, 61, 62, 63]. A descriptor based on magnetic deformation is used as a DFT-based proxy to find 30 magnetocaloric candidates, and among those MnCoP is measured to have a T_C of 575 K and a ΔS_M of 6 $\mathrm{J\,kg^{-1}K^{-1}}$ [64]. For lower temperature applications, HoB_2, with a remarkable ΔS_M of 40 $\mathrm{J\,kg^{-1}K^{-1}}$ and a T_C of 15 K is discovered through a HTS employing data mining and machine learning [29]. With the power of natural language processing, 6 Heusler compounds are identified by a mahine learning modeling trained on a database of 2,910 magnetocaloric compounds that is autogenerated by sourcing from scientific literature [65]. In a study on alloyed magnetocaloric materials, a dataset is collected on published experimental results on Fe2P-type magnetocaloric compounds and a machine learning model is used for composition optimization to find a promising alloy $Mn_{1.70}Fe_{0.30}P_{0.63}Si_{0.37}$, which shows a ΔS_M of around 10 $\mathrm{J\,kg^{-1}K^{-1}}$ in when below 100 K as confirmed by experiments [66]. HTS aids the discovery of various types of magnetocaloric materials for different target temperatures. However, limited candidates were identified so far due to limitations in the size of available MCE datasets and lacked an efficient and accurate computational methods to validate the performance of predicted materials. Chapter 3 will present a work that offers a resolution to mitigate

this obstacle. By overcoming existing challenges, MCE materials have the potential to revolutionize cooling technologies in various settings and contribute to a more environmentally friendly future.

Chapter 2

Search for Semiconductors with Ultralow Thermal Conductivity by Anti-bonding Valence Bands

2.1 Introduction

Crystalline solids with a low thermal conductivity have been extensively studied for applications such as thermoelectric materials [21, 19, 20, 14], thermal insulation and thermal barrier coatings [67]. For example, to achieve highly efficient thermoelectric power generation or cooling, one has to overcome the challenge to minimize the lattice thermal conductivity κ while maintaining the electrical conductivity σ to maximize the thermoelectric figure of merit $ZT = \frac{\sigma S^2 T}{\kappa}$, wherein S is the Seebeck coefficient and T is the temperature. A high figure of merit ZT in crystalline solids requires that the electron flow remains unimpeded while phonons get heavily scattered, a scenario known as the "phonon-glass, electron-crystal" (PGEC) [12, 13, 68, 11]. Some of the most successful techniques developed for the manipulation of phonon transport include

introducing defects [33, 15], nano/microstructural modifications using grain-boundary engineering [16, 17] and interfaces [18]. However, these extrinsic approaches all introduce defects and/or interfaces that can also interfere with electron transport [14]. In this light, crystalline materials with an intrinsic ultralow thermal conductivity are particularly desirable. In previous studies, intrinsically low thermal conductivities have been achieved by exploring phonon scattering mechanisms originating from chemical bonding and structural aspects [69, 70, 71, 72, 73]. These include layered structures [74, 75], liquid-like sublattices [76, 77, 78], local structural distortions [79], ferroelectric-instability-induced phonon softening [40, 80], rattling phonon modes [81, 82, 83, 84], and anharmonic lattice vibrations originating from electron lone pairs [85, 86, 87].

At a microscopic level, the intrinsic lattice thermal conductivity of crystalline materials is largely controlled by the chemical bonding strength, which significantly impacts the velocity of the heat-carrying acoustic phonons. This explains the much lower thermal conductivity in materials with dominant ionic or van der Waals bonds than those with dominant covalent bonds [88]. However, ionic and van der Waals bonds tend to have limited overlap between neighboring electronic orbitals, leading to weak electronic dispersion and low electron mobility. To balance the impact of the chemical bonds on the transport properties of both phonons and electrons, special types of covalent bonds with a mixed ionic character have been pursued. One prominent example is the abnormally low intrinsic lattice thermal conductivity of the IV-VI semiconductor family that hosts many of the best known thermoelectric materials, including PbTe, SnSe and GeTe [89, 90, 91, 92]. Despite their simple and high-symmetry crystal structures, these materials exhibit much lower thermal conductivity than other materials with similar atomic masses [93]. One common feature underlying their low lattice thermal conductivity is the presence of soft optical phonon modes that can strongly scatter heat-carrying low-frequency acoustic phonons, which originate from the long-ranged interatomic in-

teractions due to resonant chemical bonds [93, 94]. It is also understood that the lone electron pair associated with the group IV element in these materials plays an important role in their unusual bonding structure [85, 95, 96]. Due to their mixed ionic-covalent bonding character, the electron mobility is not compromised in these materials, making them promising platforms to realize PGEC. Another emerging example with an ultralow lattice thermal conductivity is the lead halide perovskites (LHPs), which have attracted significant recent attention for photovoltaic applications [97]. Both fully inorganic and inorganic-organic hybrid versions of these materials have a thermal conductivity below 1 W/(m K) at room temperature, which has been attributed to soft chemical bonds[98]. It has been increasingly recognized that the mixed ionic-covalent character of the chemical bonds in LHPs, likely also linked to the lone electron pair in Pb^{2+} ions [99], is not only responsible for their strong lattice anharmonicity, low lattice thermal conductivity and facile ion migration [100], but also their extraordinary optoelectronic and electron transport properties. Given the detailed understanding of the unusual lattice properties of these materials, the remaining challenge is to efficiently identify other materials with similar characters that can be promising candidates for applications requiring an intrinsically low thermal conductivity.

Indeed, the quests for new materials with extremely low thermal conductivity have been abundant. Previous studies have utilized high-throughput search techniques to identify potential candidates based on physical characteristics such as large atomic mass [23], structural information of rattling atoms [101], and the combination of machine-learning algorithms and automatic *ab initio* calculations [24, 102]. However, a comprehensive search focusing on the chemical bonding character and its impact on lattice dynamics has not been carried out due to the lack of an effective descriptor. In this work, we focus on a chemical bonding signature: highest-occupied valence bands with a strong anti-bonding character in a semiconductor. Recent studies have suggested that anti-

bonding chemical bonds are closely related to ultralow thermal conductivities in a range of materials [103, 104, 105]. The advantage of using the anti-bonding character of the highest-occupied valence band as a descriptor is that it can be efficiently analyzed using the crystal orbital Hamilton populations (COHP) method [106, 107], which only requires ground-state density functional theory (DFT) calculations. This method can be applied to any inorganic crystal structure and requires only basic structural and compositional information as input and minimal time and computing resources. The low computational cost makes it a suitable method to screen materials in databases to identify new semiconductors with an ultralow lattice thermal conductivity. In this paper, we first show that the strong anti-bonding valence bands (ABVBs) underlie the unusual lattice dynamics in known example materials with an ultralow thermal conductivity, such as PbTe and $CsPbBr_3$. In particular, the highest occupied valence bands with a strong anti-bonding character not only lead to weakened chemical bonds, but also give rise to soft optical phonons. Then, we conduct a high-throughput screen within the Materials Project database [108] to search for binary semiconductors with a strong ABVB. As a result, we identified more than 600 binary semiconductors with an anti-bonding valence band from a pool of over 6,000 candidates, allowing for the possibility of realizing ultralow intrinsic thermal conductivities. Out of the identified candidates, we highlight the XS (X=Na, K, Rb, and Cs) semiconductor family, where the sulfur ions are in an unusual valence state (S^-). We found that they all exhibit ultralow thermal conductivities despite having simple crystal structures, which can be attributed to their strongly ABVBs. Our findings suggest that a highest-occupied valence band with a strong anti-bonding character is indicative of impeded thermal transport and our high-throughput screening strategy also offers a novel approach for identifying materials with an intrinsically low thermal conductivity.

2.2 Methods

Our density functional theory (DFT) electronic structure calculations were performed using the Vienna Ab initio Simulation Package (`VASP`) [109, 110] with the projector augmented wave (PAW) method [111] and the Perdew-Burke-Ernzerhof form of the generalized gradient approximation (PBE-GGA) of the exchange-correlation functional [112]. The valence wavefunctions were expanded on a plane-wave basis with a 400 eV energy cutoff for all the materials studied in this work. The spin polarization and spin-orbit interaction were explicitly taken into account. The energy and force convergence criteria were set to be 1×10^{-7} and 0.01 eV/Å, respectively. The Γ-centered $n \times n \times n$ (n = 1, 2, 3, or 4) **k**-point grids were used to sample the first Brillouin zone depending on the unit cell size during high-throughput material screening. For selected materials studied in this paper, further **k**-point convergence was tested to make sure the lattice parameters and forces were converged. Particularly, the bonding nature associated with various electron energy bands were analyzed using the COHP method [106, 107]. The COHP method partitions the energy of the band structure into interactions between pairs of atomic orbitals between adjacent atoms. It is a bond-weighted measure of the electronic density of states (DOS) and provides a quantitative measure of the bonding and anti-bonding contributions to the band energy. Importantly for this work, the sign of the COHP differs for bonds with bonding or anti-bonding nature: a positive (negative) sign corresponds to anti-bonding (bonding) interactions. By convention, COHP diagrams plot the negative value (-pCOHP) such that bonding (anti-bonding) states on the right (left) of the axis can be easily visualized. We further quantify the degree of anti-bonding using the integrated area under the COHP curves with respect to the electron band energy.

The phonon dispersion relations were obtained by conducting finite-displacement force calculations [113], from which the harmonic interatomic force constants (IFCs)

were extracted. Then the dynamic matrices were constructed and diagonalized using the Phonopy package to generate phonon eigenfrequencies [114]. Convergence of the phonon dispersions with respect to the supercell size and the **k**-grid sampling were tested in materials selected for a detailed study. The phonon dispersion calculation of polar materials also included the non-analytic polar corrections. We calculated the lattice thermal conductivity κ_{ph} by iteratively solving the phonon Boltzmann transport equation (BTE) using the ShengBTE package [115]. Using the finite displacement method, we computed the anharmonic 3rd-order IFCs. The 3rd-order finite displacement calculation employed the identical supercell size and **k**-grid sampling as the 2nd-order IFCs. To ensure convergence, we tested several neighbor interaction cutoffs. The q-mesh density, which is the phonon momentum space sampling grid, was set to $10 \times 10 \times 10$ for most materials. We examined the convergence of grid density for all cases, and the reported values of thermal conductivity are all converged.

2.3 Results and Discussion

2.3.1 Anti-bonding valence bands in PbTe and $CsPbBr_3$

In the chemical bonding theory [116, 117], molecular orbitals (MOs) are formed by the overlap of atomic orbitals. For example, Fig. 3.1(a) illustrates the π bond formation of atomic p orbitals between two adjacent atoms. The overlap of electron wave functions splits the original atomic p orbital into two MOs. The lower energy orbital is the bonding MO, shown as the "bonding π" in Fig. 3.1(a). The higher energy orbital is the anti-bonding MO, shown as the "anti-bonding π^*". Due to the Pauli exclusion principle, the two electrons will occupy the two spin states of the bonding MO, leaving the anti-bonding MO unoccupied. Consider the energy level of the two MOs as a function of

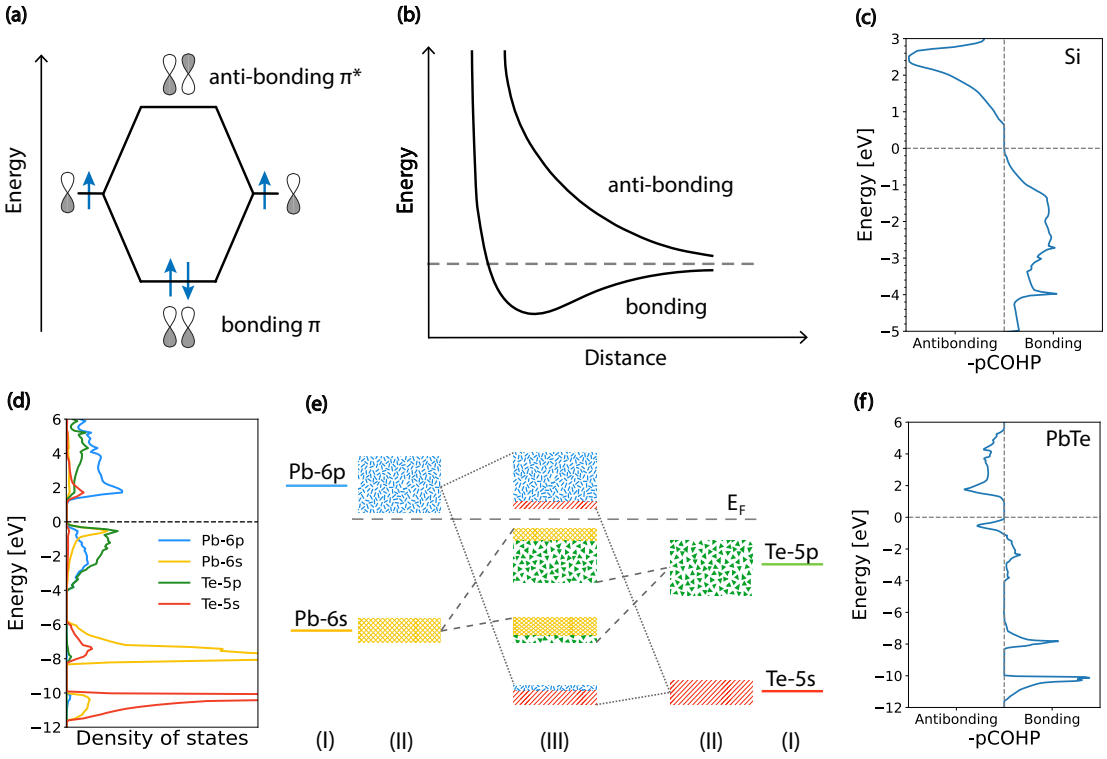

Figure 2.1: **Schematics of anti-bonding states and anti-bonding valence bands.** (a) Molecular orbital diagram of a π bond formed by p orbitals. (b) Energy of typical bonding and anti-bonding orbitals as a function of atomic distances. (c) COHP diagram of silicon, showing the bonding nature of the valence band and the anti-bonding nature of the conduction band. (d) Electronic density of states of PbTe, projected onto atomic orbitals. (e) Molecular orbital diagram of PbTe, showing three stages of orbital interactions: (I) isolated Pb and Te atomic levels; (II) broadened atomic levels due to s-s and p-p mixing; (III) s-p hybridization introduces s orbital features to valence band maximum (VBM) and conduction band minimum (CBM) and an occupied anti-bonding valence band below the Fermi level E_F. (f) COHP diagram of PbTe, showing the strong anti-bonding nature of the valence band.

the atomic distance, as shown in Fig. 3.1(b). The occupation of the bonding MO lowers the total energy near the equilibrium interatomic distance and results in stabilization of the chemical bond, while the occupation of the anti-bonding MO raises the energy and, thus, weakens the bond and potentially increases the bond anharmonicity. Many of the covalently bonded semiconductors possess a valence band with occupied bonding states and a conduction band with empty anti-bonding states. For example, the valence band and the conduction band in silicon are formed by the bonding and anti-bonding states of the sp^3 hybridized orbitals, respectively. This can be clearly seen from the COHP analysis of silicon, which is shown in Fig. 3.1(c).

In contrast, the bonding nature of the electronic bands in PbTe is qualitatively different. PbTe has been extensively studied due to its exceptional thermoelectric properties including a high Seebeck coefficient and a low thermal conductivity [33, 39, 42]. Despite its high-symmetry structure, the bulk lattice thermal conductivity in single-crystalline PbTe is unexpectedly low, around $2\,\mathrm{W/(m\ K)}$ at room temperature [42]. Fig. 3.1(d) shows the PbTe electronic density of states (DOS) that has been projected onto Pb-6p, Pb-6s, Te-5p, and Te-5s orbitals. Based on the electronic DOS, Fig. 3.1(e) shows a simplified linear combination of atomic orbitals (LCAO) picture of PbTe. Pb^{2+} ions have empty 6p states, but the 6s states are occupied by a "lone pair" of electrons [118]. This unique "lone pair" configuration enables efficient s-p mixing between the Pb-6s states and the Te-5p states, forming bonding and anti-bonding states. Since both Pb-6s and Te-5p states are fully occupied, both the resulted bonding and anti-bonding states are occupied, leading to the highest-occupied valence band possessing a strong anti-bonding character, as shown in the COHP analysis shown in Fig. 3.1(f). This intimate relationship between the lone-pair electrons of the divalent group-IV cations and the ABVB was also discussed in previous studies [118]. However, a direct evaluation of the impact of the ABVB on the thermal transport is still lacking.

To quantify the ABVB effect on thermal transport properties in PbTe, we hypothesized that, by removing electrons from the anti-bonding states at the VBM and then relaxing the lattice, the chemical bond strength will be increased and the bond anharmonicity will be reduced, leading to a stabilized lattice and an increased thermal conductivity. To test this hypothesis, we used a $4 \times 4 \times 4$ supercell in the calculation, where electrons were removed in PbTe and the lattice was fully relaxed afterwards. We performed the calculation by removing 2 and 8 electrons (labeled "-2e" and "-8e" cases, respectively, in the following discussion) out of a total number of 640 valence electrons, which can be considered a small perturbation to the pristine PbTe electronic structure and only represent the anti-bonding states very close to the VBM. A 4-fold degeneracy exists for 8 VBM electrons located at L (0.5, 0.5, 0.5) points, which are shifted to the Γ point in the supercell calculation. Therefore, a single Γ point sampling was sufficient in the $4 \times 4 \times 4$ supercell.

The calculated IFCs and the bond lengths are shown in Table 1, where we compare three cases: unperturbed PbTe, the "-2e" case and the "-8e" case. Here, K_{12} is the trace of the IFC tensor of the nearest Pb-Te bond with its definition explained in Sec. 2.5.1. The negative sign of the K_{12} values indicates the overall stable bonding character of the Pb-Te bond and their magnitude reflects the bonding strength. From Table 1, it is clearly seen that removing electrons from the ABVB states decreases the length of the Pb-Te bond and increases its bonding strength, thus stabilizing the PbTe lattice. This effect is further illustrated in Fig. 3.2(a) as the total ground-state energy of the unperturbed PbTe and the "-2e" case is plotted as a function of the lattice constant. The minimum location, corresponding to the equilibrium lattice parameter, shifts towards a smaller value in the "-2e" case, with an increased quadratic coefficient from 0.0304 to 0.0319, indicating a strengthened bond. The difference in the ground-state energy between the two cases, ΔE, reflects the contribution from the anti-bonding component, showing a

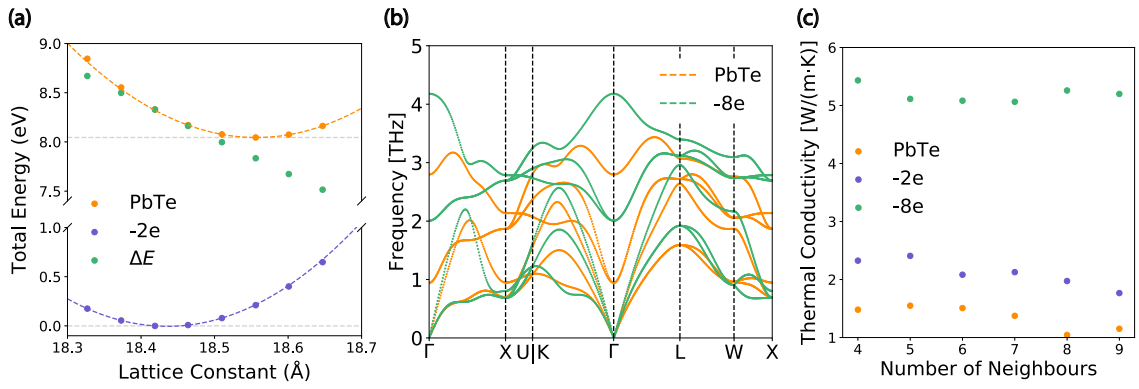

Figure 2.2: **Impact of the anti-bonding valence band on thermal transport in PbTe.** (a) Total ground-state energy of unperturbed PbTe and the "-2e" case as a function of the supercell lattice constant. The difference of the two energies, ΔE is also shown, which reflects the contribution of the removed anti-bonding states near the VBM. The energy axis is shifted such that the energy of the "-2e" case at the equilibrium volume is zero. (b) Calculated phonon dispersion relations of the unperturbed PbTe and the "-8e" case, showing the impact of anti-bonding states on the acoustic phonon velocity and the soft optical phonon frequency.(c) Calculated lattice thermal conductivity of the unperturbed PbTe and the "-2e" and "-8e" cases as a function of the interacting neighbor shells included in the calculation, showing the significant impact of the occupied anti-bonding valence states on the thermal conductivity of PbTe.

destabilizing trend.

Table 1: Traces of the IFC Tensor and Bond Length of PbTe

IFC	PbTe	-2e	-8e
K_{12}	-0.786	-0.869	-1.18
Bond Length	3.28Å	3.26Å	3.19Å

The impact of removing electrons from ABVB states on phonon properties is examined in Fig. 3.2(b), which illustrates the phonon dispersions of the unperturbed PbTe and the "-8e" case. Firstly, the acoustic phonon modes have an increased velocity in the "-8e" case compared to that in the unperturbed case, which is a consequence of the increased bonding strength consistent with the magnitude of K_{12} in Table 1. Secondly, removing electrons from the ABVB states significantly raises the frequency of the soft

optical phonons near the Γ point, suggesting that the occupied ABVB in PbTe not only weakens the chemical bond, but also leads to long-ranged force interactions resulting in soft optical phonons. Both effects are expected to strongly affect the thermal conductivity. Fig. 3.2(c) shows the calculated thermal conductivity in unperturbed PbTe as compared to "-2e" and "-8e" cases as a function of the number of neighbor shells included in calculating the third-order IFCs. The converged PbTe thermal conductivity is close to the experimental value [42] and previous first-principles calculation [119]. The corresponding scattering properties of PbTe are included in Sec. 2.5.2. The thermal conductivity in the "-2e" case is approximately 1.5 times that of the unperturbed PbTe, while in the "-8e" case, the thermal conductivity is boosted by at least a factor of 3. These findings indicate that the presence of occupied ABVBs is the origin for the abnormally low lattice thermal conductivity in PbTe from a chemical-bonding-theory point of view. To solidify these findings, we further applied the same strategy (removing valence electrons) to Si with bonding valence bands, and the results can be found in Sec. 2.5.1. In contrast to PbTe with ABVB, removing valence electrons in the bonding states in Si leads to a weakened bond strength, a decrease in the speed of sound and a reduction in the thermal conductivity [120]. This result further establishes the impact of ABVB on the intrinsic lattice thermal conductivity.

Similarly, we also examined the relationship between the ABVB and the thermal transport properties of the inorganic halide perovskite $CsPbBr_3$ that also contains the Pb^{2+} ion. Halide perovskites have gained attention as a new class of photovoltaic and thermoelectric material [121, 122, 97, 123]. While the ultralow thermal conductivity of organic-inorganic hybrid halide perovskites was often attributed to the cation dynamic disorder, it is not well understood why thermal transport in crystalline all-inorganic halide perovskites such as $CsPbBr_3$ is also significantly suppressed [124, 79]. Similar to PbTe, the Pb^{2+} ion in $CsPbBr_3$ hosts Pb-6s lone pair electrons [125, 126, 127] that promote

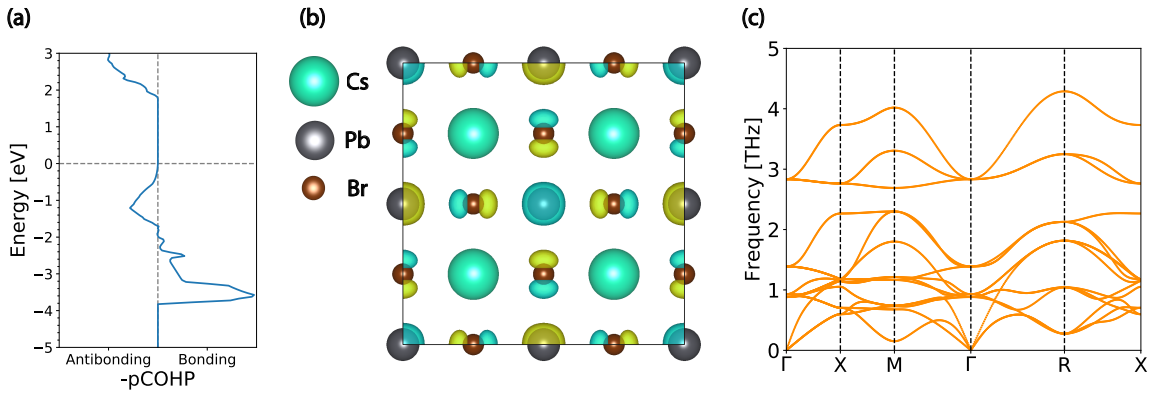

Figure 2.3: **Anti-bonding valence bands in $CsPbBr_3$ and their impact on phonons.** (a) The COHP diagram for $CsPbBr_3$, where a strong anti-bonding feature is shown for the VBM. (b) Isosurfaces of the electronics wavefunctions at the VBM in $CsPbBr_3$. The s orbitals on Pb atoms and p orbitals on Br atoms show anti-bonding interactions as reflected by the opposite signs of the wavefunctions associated with facing Pb and Br atomic pairs. (c) The calculated phonon dispersion of the cubic phase of $CsPbBr_3$ at 400 K. Low acoustic velocity associated with the anti-bonding valence bands is observed, which drives the low intrinsic thermal conductivity.

the s-p mixing between Pb-6s and Br-4p orbitals, leading to an ABVB in $CsPbBr_3$ that can be quantified by the COHP analysis shown in Fig. 3.3(a). A visualization of the Pb-Br anti-bonding interaction is shown in Fig. 3.3(b), where isosurfaces of the electronic wavefunctions at the VBM of $CsPbBr_3$ are shown. Here, blue and green isosurfaces indicate opposite signs of the wavefunction and the opposite signs of the isosurfaces associated with facing Pb and Br pairs confirm the anti-bonding nature of the valence band in $CsPbBr_3$. Figure 3.3(c) shows the phonon dispersion relation of the cubic phase $CsPbBr_3$ at 400 K simulated by the temperature-dependent effective potential (TDEP) method [128], where the low acoustic phonon velocity originated from the weakened bonds due to the ABVBs is clearly seen that drive the low lattice thermal conductivity. Figure 3.3(c) is also similar to the phonon dispersions in the literature [129, 130].

2.3.2 High-throughput Screening of Semiconductors with ABVBs

Results from PbTe and $CsPbBr_3$ suggest that a strong anti-bonding character in the highest occupied valence band, which can be efficiently evaluated by the COHP analysis, is a convenient indicator for weakened covalent bonds, large phonon anharmonicity and a low lattice thermal conductivity. This motivates us to perform a high-throughput screening for semiconductors with strong ABVBs in the Materials Project database [108] by the COHP analysis. As a first step, we focused on binary compounds and found ∼1000 binary candidates in the Materials Project database that are stable and have a finite band gap between 0 to 3 eV. Materials containing heavy elements with unfilled f electrons were excluded from the candidate list due to the known inaccuracy of DFT to deal with these materials. For each candidate, we calculated COHP for all pair-wise interactions within 1.5 times the nearest-neighbor distance. Then the strength of the anti-bonding character was quantified by integrating the area under the COHP curves within 0.1 eV of the valence band edge, thus only focusing on the highest-occupied valence band states. All candidates were sorted by the strength of the anti-bonding character associated with their highest occupied valence band states. A screened list consisting of 625 experimentally stable binary semiconductors are provided in Sec. III of the Supplementary Material in [1]. Among these materials of interest, we present a comprehensive study of the thermal transport of the XS family, where X = Na, K, Rb, Cs are alkaline metals. Their structural properties and thermal conductivity at room temperature are summarized in Table 2, and the convergence tests of their thermal conductivities are provided in Sec. 2.5.4. As shown in Table 2, with an increasing mass of the alkaline metal, the room-temperature thermal conductivity decreases from 5 W/(m

[1] https://pubs.acs.org/doi/suppl/10.1021/jacs.3c05091/suppl_file/ja3c05091_si_001.pdf

K) in NaS to 0.1 W/(m K) in CsS. KS, RbS, CsS all show an ultralow room-temperature thermal conductivity < 0.8 W/(m K) despite their relatively simple crystal structures. A comparison between the thermal conductivities calculated using the full iterative solution of the phonon BTE and the relaxation time approximation (RTA) is included in Table 5 in Sec. 2.5.5. The calculated thermal conductivities using the two methods are similar, suggesting the dominant role played by Umklapp processes in phonon-phonon scattering due to their soft and strongly anharmonic lattices. Given their moderate bandgaps above 1 eV as shown in Table 2, their intrinsic electronic thermal conductivity should be negligible given the low intrinsic concentration of charge carriers. The results open up possibilities for their potential applications in thermoelectrics and also motivate further experimental studies. Notably, KS reaches an ultralow thermal conductivity of 0.8 W/(m K) at room temperature without heavy elements or complex structures. A detailed analysis of the bonding mechanism and the thermal transport was thus carried out for KS as an example to demonstrate the decisive impact of the ABVB on its ultralow thermal conductivity.

Table 2: Basic properties and calculated thermal conductivity of NaS, KS, RbS, CsS

Material	Band Gap [eV]	Crystal System	Space Group	κ_x	κ_y	κ_z	[W/(m K)]
NaS	1.23	hexagonal	$P6_3/mmc$	5.9	5.9	3.9	
KS	1.47	hexagonal	$P\bar{6}2m$	0.84	0.84	0.76	
RbS	1.58	hexagonal	$P\bar{6}2m$	0.65	0.65	0.64	
CsS	1.73	orthorhombic	Immm	0.15	0.14	0.083	

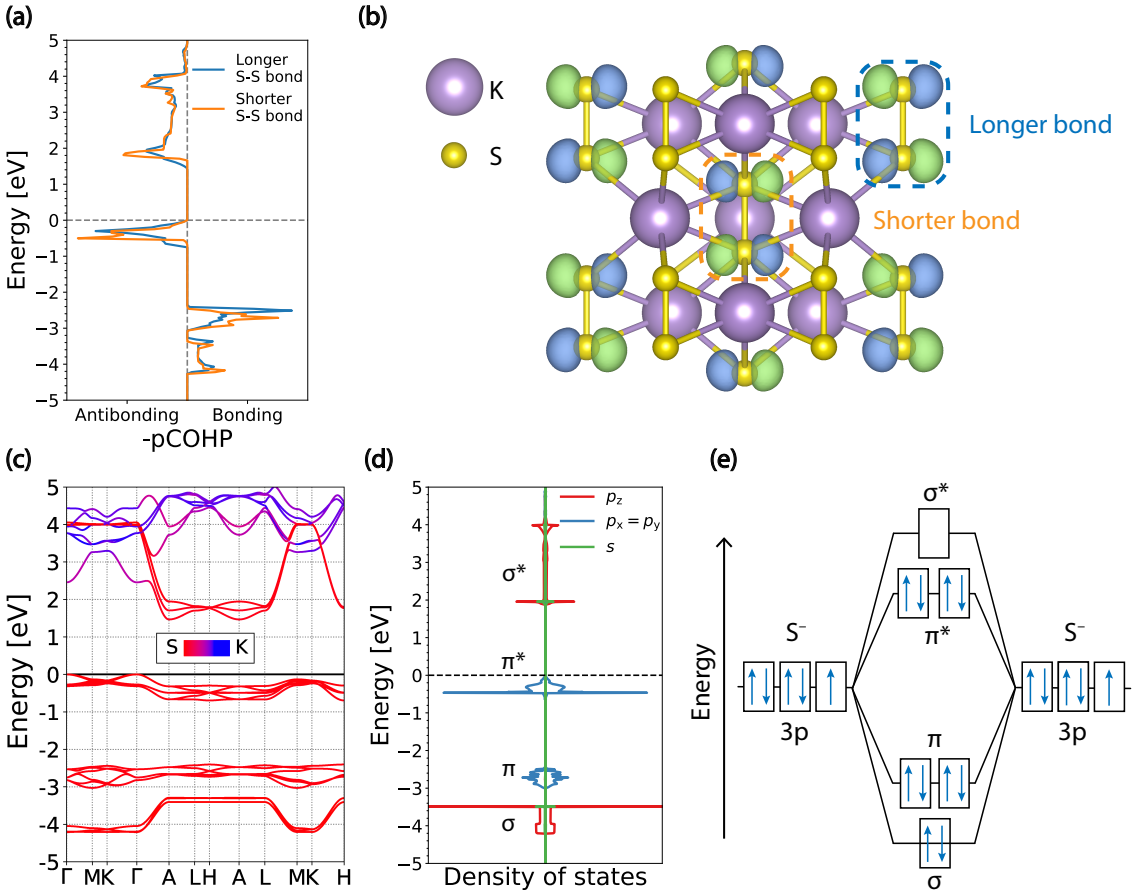

Figure 2.4: **Anti-bonding valence bands in KS.** (a) The COHP diagram of KS, where the S-S bonds show strong anti-bonding features near the VBM. The longer and shorter S-S bonds correspond to a bond length of 2.15 Å and 2.13 Å, respectively. (b) The isosurfaces of the VBM electronic wavefunctions in KS. The longer and shorter S-S bonds are marked in blue and orange boxes, respectively. (c) The calculated electronic band structure of KS. The S orbitals contribute dominantly to the valence band. (d) l-decomposed and site-projected electronic DOS for S orbitals in KS. p_x (equivalent to p_y) orbitals form the valence bands, and the highest occupied valence band consists of the anti-bonding π^* state formed by p_x (p_y) orbitals. (e) Molecular bond analysis for the S-S bond. The 3p orbitals of a single S^- ion has 5 electrons. The covalent S-S bond form σ, σ^*, π and π^* orbitals, where σ, π and π^* bonds are occupied. These occupied bonds correspond to the electronic DOS below the Fermi level as shown in Fig. 2.4(d) and the highest occupied valence band is composed of π^* anti-bonding states.

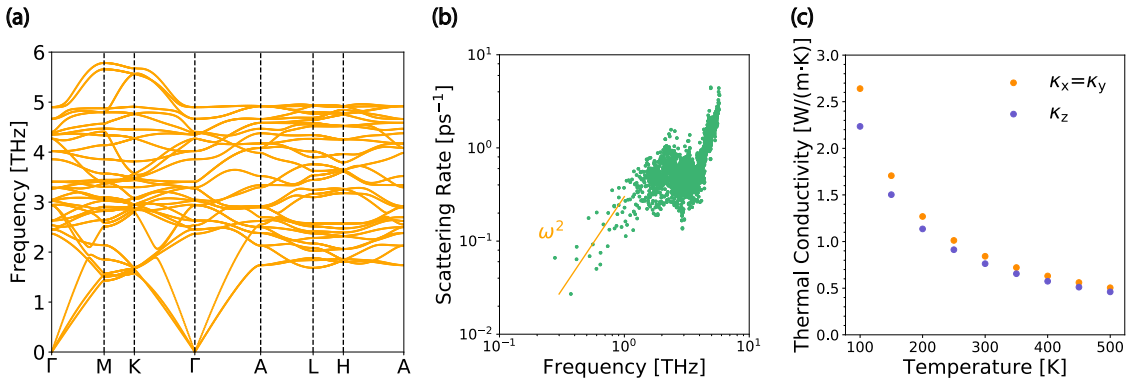

Figure 2.5: **Calculated thermal transport properties of KS.** (a) The calculated phonon dispersion relation of KS, which shows the material is dynamically stable. Flat phonon bands are also observed near 1.5 THz. (b) The calculated phonon-phonon scattering rate in KS at 300 K. The scattering rates of the low-frequency acoustic phonons follow the frequency-squared scaling, while flat phonon bands lead to increased scattering near 1.5 THz. (c) The calculated temperature-dependent lattice thermal conductivity of KS from 100 K to 500 K, showing an ultralow lattice thermal conductivity below 1 W/(m K) at room temperature.

2.3.3 Detailed Analysis of Bonding and Thermal Transport in KS

Despite its unusual chemical formula, KS (or K_2S_2) has been theoretically found to be as stable as K_2S at room temperature under ambient pressure [131] and has been experimentally synthesized [132, 133]. Building upon these results, we conducted a more in-depth analysis of KS, focusing specifically on its chemical bonding properties. The COHP diagram shown in Fig. 2.4(a) attributes the strong anti-bonding character of the valence band to two kinds of S-S bonds. These two S-S bonds can be visualized in the isosurfaces of VBM electron wavefunctions as shown in Fig. 2.4(b), where the blue and green isosurfaces represent opposite signs of the wavefunction. From the isosurfaces, the two anti-bonding S-S bonds are π^* bonds of the S-3p orbitals, where the neighboring wavefunctions possess opposite signs. The longer bond (2.15 Å), marked in a blue box in Fig. 2.4(b), has more prominent anti-bonding component closer to the VBM. The shorter

bond labeled in an orange box shows stronger anti-bonding due to its shorter (2.13 Å) length. The calculated electronic band structure in Fig. 2.4(c) shows that orbitals from the S atom make dominant contributions to electronic states near the band edges, while K orbitals only contribute to unoccupied higher energy conduction bands. The effective masses of electrons and holes near the band edges can also be extracted from the calculated band structure: the electron effective mass is $0.24\,m_0$ along the out-of-plane direction and $1.48\,m_0$ along the in-plane direction; the hole effective mass is $1.2\,m_0$ along the out-of-plane direction and $3.4\,m_0$ along the in-plane direction, where m_0 is the free electron mass. The electronic band structures and effective masses of the other members of the family are provided in the Supplementary Material Sec. II.C. All members have relatively low effective electron and hole masses. The relatively low effective masses and the moderate bandgap show that the electronic transport properties are maintained despite the reduced bonding strength leading to a lower thermal conductivity and suggest that ABVB materials can possess a potentially high intrinsic electron mobility and, if properly doped, can reach a high electrical conductivity. This is in contrast to ionic crystals, whose usually large bandgaps and heavy effective masses preclude them from being effectively doped to reach a high electrical conductivity. From the orbital-decomposed and site-projected electronic DOS for an S atom shown in Fig. 2.4(d), we can observe that the equivalent p_x and p_y orbitals form the anti-bonding π^* state that corresponds to the highest occupied valence band in KS. The molecular orbital diagram of the S-S bond in KS is illustrated in Fig. 2.4(e) to reveal the origin of the ABVB in KS. The outermost 4s electron in a K atom is transferred to an S atom to form a S^- ion in its unusual monovalent configuration. The 3p orbitals of a single S^- ion thus contain 5 electrons. The covalent S-S bond form σ, σ^*, π and π^* orbitals, where σ, π and π^* orbitals are occupied, yielding a bond order of 1. Here the anti-bonding π^* state is occupied thanks to the additional electron transferred from the K atom. These occupied orbitals

correspond to the electronic DOS below the Fermi level in Fig. 2.4(e). Therefore, the highest occupied valence band is composed of anti-bonding π^* states. This mechanisms is in addition to the group-IV lone-pair electrons that can give rise to ABVBs. Recently, He et al. also pointed out that p-d hybridization can lead to ABVBs and ultralow thermal conductivities in Cu- and Ag-containing compounds [73]. These chemical insights into ABVB formation can further guide the search and design of new materials with an ultralow intrinsic thermal conductivity.

To complete our discussion, we report the calculated thermal transport properties of KS in Fig. 2.5. Figure 2.5(a) shows the phonon dispersion, suggesting that KS is dynamically stable. More importantly, flat phonon bands can be observed near 1.5 THz, which are expected to scatter acoustic phonons and hinder thermal transport. Figure 2.5(b) depicts the phonon-phonon scattering rates as a function of the phonon frequency, which are compared to the expected frequency-squared scaling. The scattering rates of the low-frequency acoustic phonons follow the frequency-squared trend while the flat phonon bands lead to a peak in the scattering rates near 1.5 THz. The calculated lattice thermal conductivity of KS is 0.8 W/(m K) at room temperature, which is lower than most compounds with similar atomic mass and simple crystals structure. The temperature-dependent thermal conductivity of KS from 100 K to 500 K is displayed in Figure 2.5(c). The compound CsS in the same family with a heavier alkaline metal even possesses a room-temperature thermal conductivity as low as 0.1 W/(m K) (more details are provided in Sec even lower than many amorphous and organic materials. These materials are promising candidates for applications where an ultralow thermal conductivity is desired.

We note that our current discussion is limited to identifying materials with an intrinsically low thermal conductivity, while we have indicated that materials with ABVBs have the potential to reach a high electrical conductivity. A complete analysis of dopa-

bility and electrical transport, however, requires a different set of calculations, including understanding the defect formation physics, determining scattering rates of electrons and solving the electron Boltzmann transport equation. Such considerations are beyond the scope of the current work and will be pursued in our future studies.

2.4 Conclusion

In conclusion, we examined the connection between ABVBs and low thermal conductivities with first-principles calculations in PbTe and $CsPbBr_3$ and we found that the ABVBs are responsible for the weakened chemical bonds and the soft optical phonons that lead to abnormally low lattice thermal conductivities in both compounds. Based on this observation, we conducted a high throughput materials search based on ABVBs and found over 600 experimentally stable binary semiconductors with strong ABVBs. Among the candidate materials, we analyzed in detail the XS (X = Na, K, Rb, and Cs) family in terms of their strong ABVB states and evaluated their lattice thermal conductivities. With an ultralow thermal conductivity of 0.1 W/(m K), CsS is an exceptional example of a crystalline material with a lower thermal conductivity than most amorphous and organic materials. Several other materials on our list with strong ABVBs are potentially interesting as well, such as $NaCl_3$, and further studies are needed to fully evaluate the impact of the ABVBs on their electrical and thermoelectric transport properties. Our material search can also be expanded to more complicated materials containing three or more elements. We note that materials with intrinsically low thermal conductivity can find other applications beyond thermoelectrics, such as thermal barrier coatings and thermal insulation. While the high strength and high melting point required by thermal barrier coatings may not be fulfilled by ABVB materials due to their generally soft lattice, the ABVB materials should be readily applicable in thermal insulation applications at

room temperature or lower temperatures. Our results suggest that the intuitions offered by the chemical bonding theory can provide simple but powerful guidelines for understanding the thermal conductivity of materials as well as for discovering new materials with unusual thermal transport properties.

2.5 Additional Data

2.5.1 Additional data for removing valence electrons in Si

To corroborate the relation between ABVB and low thermal conductivity, we employed Si with bonding valence bands (BVBs) as a comparison to PbTe with ABVB. With the convergence test for Si, we adopted a k-grid of $9 \times 9 \times 9$ and a cutoff energy of 550 eV in the 2-atom primitive cell, and the resulting lattice parameter is 3.867Å. The corresponding lattice parameter of the conventional cell is 5.466Å. The VBM of Si is located at the Γ point, and we used a single Γ-point sampling for the supercell calculations. Using a $4 \times 4 \times 4$ supercell, 6 out of 512 valence electrons were removed considering the three-fold VBM degeneracy and the spin degeneracy.

Since the VBM in Si consists of bonding states, we expect the removal of bonding electrons to weaken the chemical bond. Table 3 reports the effects of removing the valence electrons from Si. Here "Si" labels unperturbed Si, and "-6e" labels the case where 6 electrons are removed. The bond strength is weakened as K_{12} is reduced in the "-6e" case, as expected. The removed electrons case "-6e" also has a lower thermal conductivity, which means the BVB contributes to stabilizing the lattice and increasing the thermal conductivity. We further show the phonon dispersion comparison in Fig. 2.6(a), and the slope of acoustic phonons is decreased in the "-6e" case, corresponding to a decrease in the acoustic phonon group velocity. Fig. 2.6(b) shows the scattering rate of the two cases of

Si, exhibiting increased scattering rate for the "-6e" case, especially for the low-frequency acoustic phonons.

IFC tensor in Table 3 and main text Table 1 is defined as $K_{ij} = \frac{\partial^2 V}{\partial x_i \partial x_j}$, where V is the lattice energy, and x_i and x_j are coordinates of two atoms. With this definition, negative IFCs represent stable "spring" interactions between atoms while positive IFCs represent unstable "anti-spring" interactions. This convention is consistent with literature, e.g. [93]. This can be more explicitly illustrated by considering the simplest model of a lattice, where the nearest atoms are connected by stable springs. In this model, the spring interaction is represented by $V = a(x_1 - x_2)^2 + b(x_1 - x_3)^2 + \cdots$ and so forth, with the coefficients a, $b \cdots$ being positive. If we expand this interaction term, we get $V = (a+b+\cdots)x_1^2 + (a+\cdots)x_2^2 + (-2a)x_1x_2 + \cdots$. The K_{12} in Table 1 refers to the coefficient of the term x_1x_2 evaluated by $\frac{\partial^2 V}{\partial x_1 \partial x_2} = -2a$. Thus, a negative K_{12} represents a stable "spring" interaction. If K_{12} is more negative, the spring constant a is larger, and the bond becomes more stable.

Table 3: Traces of the IFC tensor, bond length, and thermal conductivity of Si with bonding valence bands (BVB). The difference between the unperturbed "Si" and the removed six electrons "-6e" cases shows that the bonding strength K_{12} decreases, and the thermal conductivity is reduced. The trend is opposite to PbTe with anti-bonding valcence bands (ABVB), and hence validates the correlation between ABVB and low thermal conductivity in crystalline materials.

IFC	Si	-6e
K_{12}	-9.81	-9.25
Bond Length	2.438 Å	2.342 Å
Thermal Conductivity	128.6 W/(m K)	97.6 W/(m K)

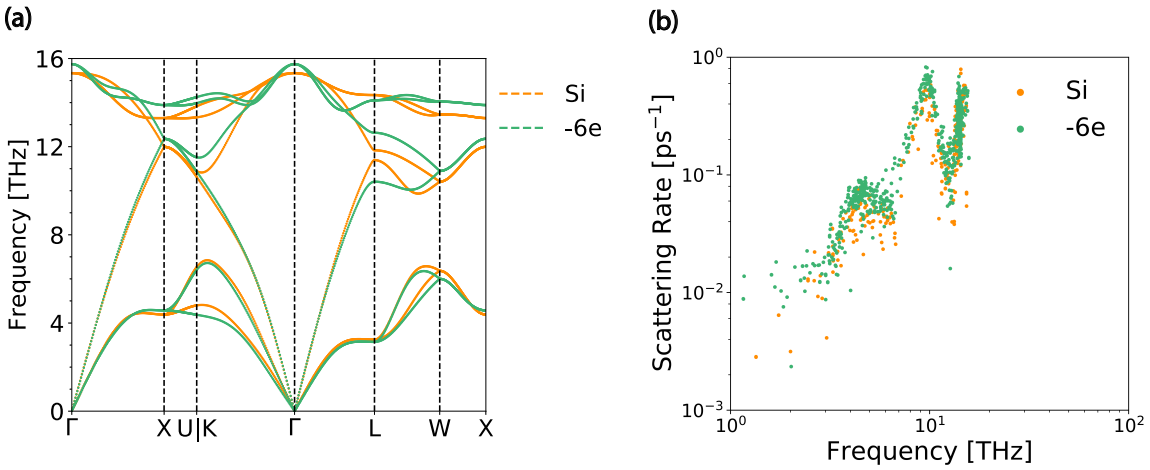

Figure 2.6: **Phonon Dispersion and Scattering Rate of Si.** (a) Phonon dispersion of Si in the unperturbed case and the "-6e" case. The slope along Γ-L and Γ-K directions was decreased in the "-6e" case compared to the unperturbed case. (b) Scattering rate of Si in the unperturbed case and the "-6e" case. The "-6e" case shows slightly higher scattering rate, especially for low-frequency acoustic phonons.

2.5.2 Additional Data for PbTe

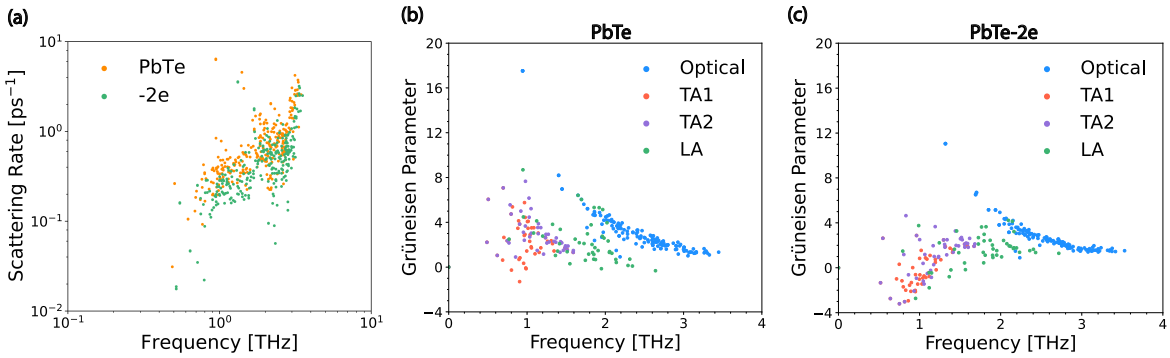

Figure 2.7: **Grüneisen Parameters and Phonon Scattering Rates of PbTe.** (a) Phonon Scattering Rate of PbTe in both the unperturbed case and the "-2e" case. The "-2e" case shows lower phonon scattering rate than the unperturbed state due to reduced anharmonicity after removing anti-bonding valence electrons. (b) Grüneisen parameters of the unperturbed PbTe and (c) Grüneisen parameters of the "-2e" PbTe show that the ABVB contributes to stronger phonon anharmonicity in PbTe, which leads to more phonon-phonon scattering and a lower thermal conductivity.

2.5.3 COHP diagrams of the XS family

Besides KS described in the main text, here we also report the COHP analysis of NaS, RbS, and CsS in the XS family shown in Fig. 2.8. All XS materials exhibit prominent anti-bonding characters near the VBM. Note that RbS has the same crystal symmetry as KS, and there are two kinds of S-S bonds: longer and shorter S-S bonds.

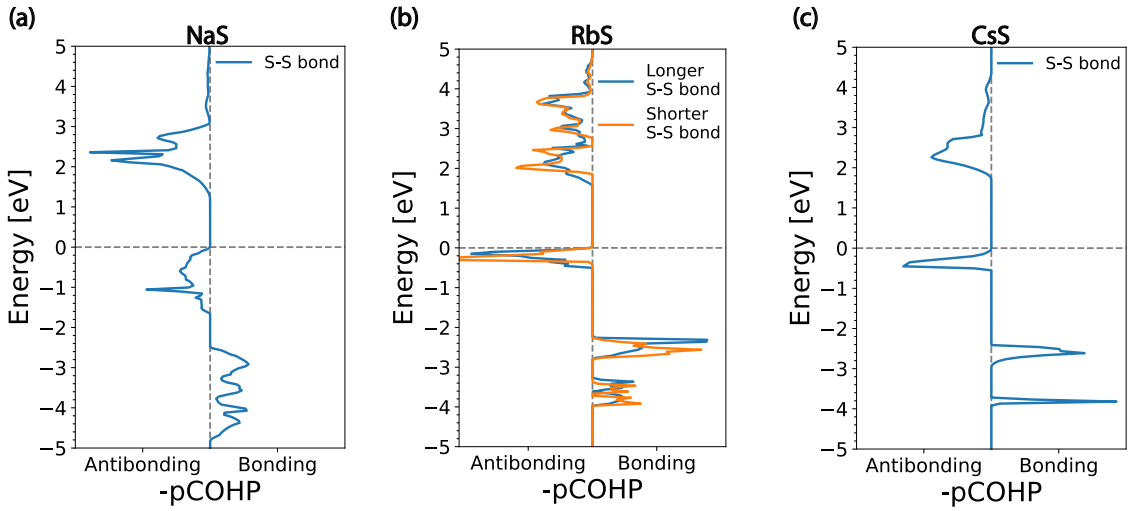

Figure 2.8: **COHP diagrams of XS(X=Na, Rb, Cs) faimily.** (a) COHP diagram of NaS displays a brodened range of anti-bonding contributions near the top of VBM. (b) COHP diagram of RbS. The VBM shows similar strong anti-bonding features as KS. (c) COHP diagram of CsS reveals that the VBM also exhibits strong anti-bonding characters.

2.5.4 Thermal properties of the XS family

We tested the convergence of thermal conductivity with the number of neighbours included in the calculation of the 3rd-order force constants for the XS materials. The change in thermal conductivity is shown in Fig. 2.9. We adopted 7, 7, 8, and 4 neighbours for NaS, KS, RbS, and CsS, respectively.

Temperature-dependent thermal conductivity from 100 K to 500 K for NaS, RbS, and CsS are shown in Fig. 2.10. The trend of thermal conductivity is consistent with the atomic weight. For CsS, which has the lowest thermal conductivity among all XS materials, its thermal conductivity is < 0.5 W/(m K) at 100 K and < 0.2 W/(m K) at 300 K. The scattering rates shown in Fig. 2.11 demonstrate a notable increase throughout NaS, RbS, and CsS. Fig. 2.12 displays the mode-specific Grüneisen parameter for all XS materials. Stronger phonon anharmonicity is shown in heavier materials for both acoustic and optical phonons.

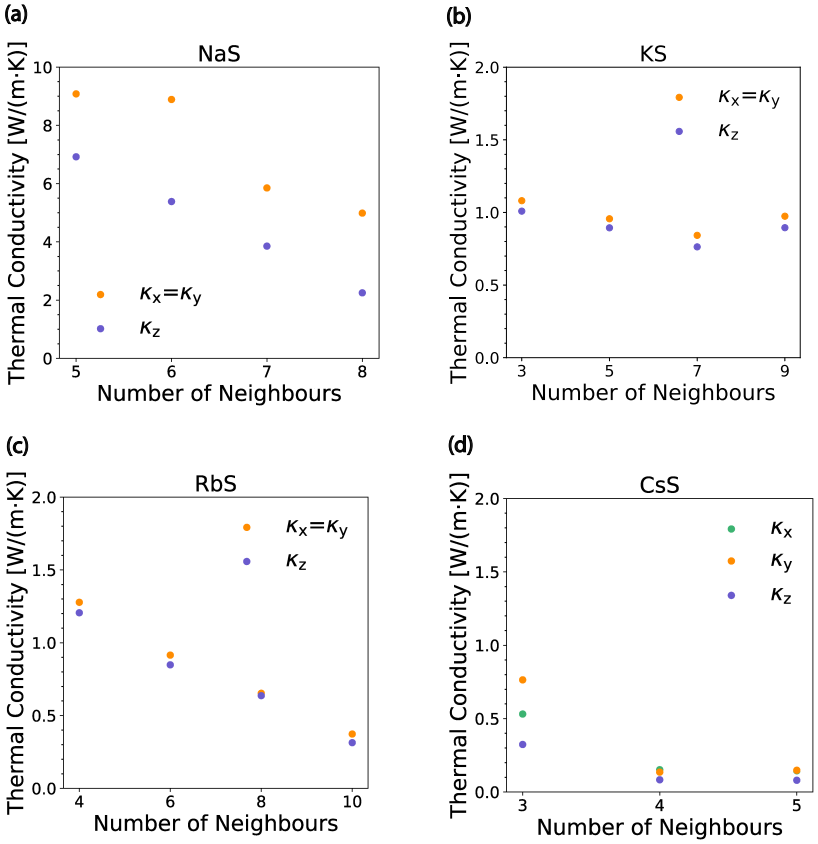

Figure 2.9: **Thermal Conductivity at 300 K as a Function of the Number of Neighbours included in the Calculation of 3rd-order Force Constants.** (a) Thermal conductivity convergence test for NaS. (b) Thermal conductivity convergence test for KS. (c) Thermal conductivity convergence test for RbS. (d) Thermal conductivity convergence test for CsS. There is a decreasing trend of thermal conductivity as the atomic mass increases in XS.

Table 4: Comparison of Relaxation Time Approximation (RTA) with iterative solution (Iterative) of 300 K thermal conductivity for PbTe and XS (X=Na, K, Rb, Cs) family.

$\kappa_x, \kappa_y, \kappa_z$	PbTe	NaS	KS	RbS	CsS
Iterative	1.374,1.374,1.374	5.85,5.85,3.85	0.842,0.842,0.763	0.653,0.653,0.637	0.152,0.135,0.083
RTA	1.005,1.005,1.005	5.62,5.62,3.93	0.841,0.841,0.798	0.649,0.649,0.621	0.114,0.125,0.068

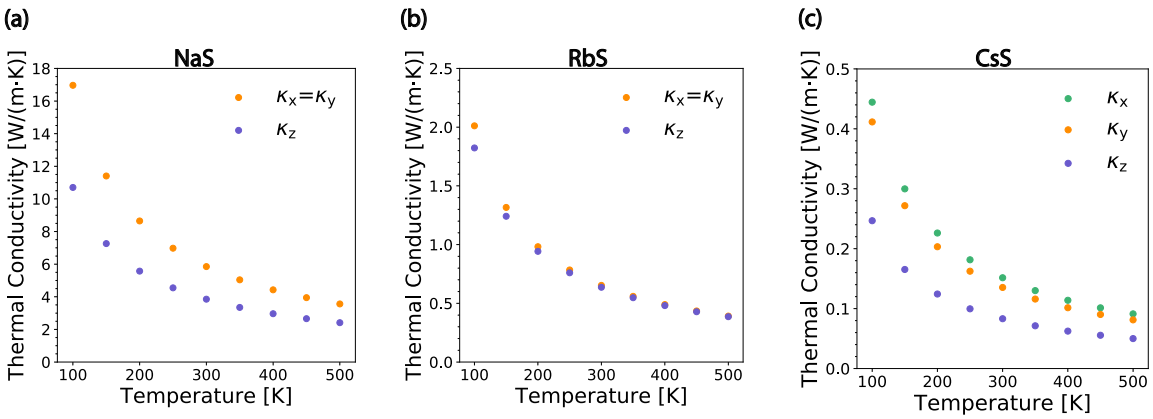

Figure 2.10: **Temperature-dependent Thermal Conductivity for NaS, RbS, and CsS from 100 K to 500 K.** (a) Temperature-dependent thermal conductivity for NaS. (b) Temperature-dependent thermal conductivity for RbS. (c) Temperature-dependent thermal conductivity for CsS.

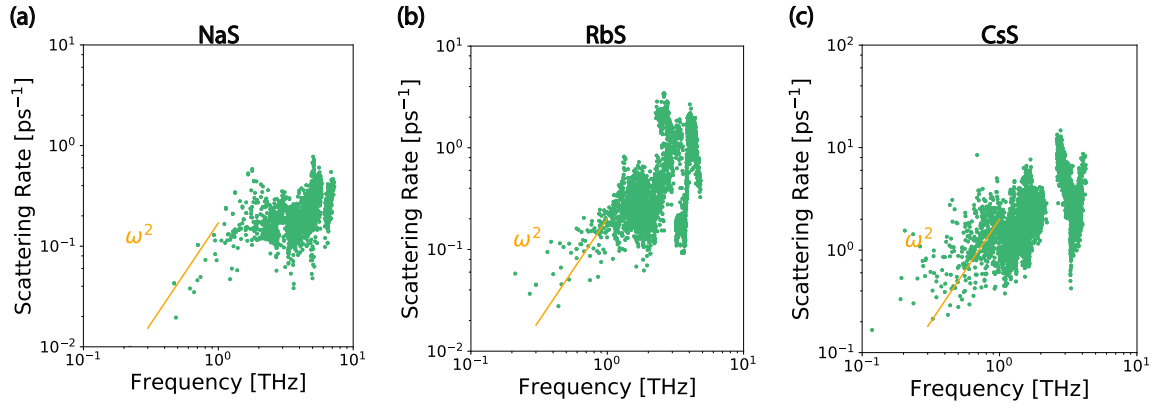

Figure 2.11: **Phonon Scattering Rates for NaS, RbS, and CsS at 300 K.** (a) Phonon scattering rates for NaS. (b) Phonon scattering rates for RbS. (c) Phonon scattering rates for CsS. The orange ω^2 lines are guides to the eye.

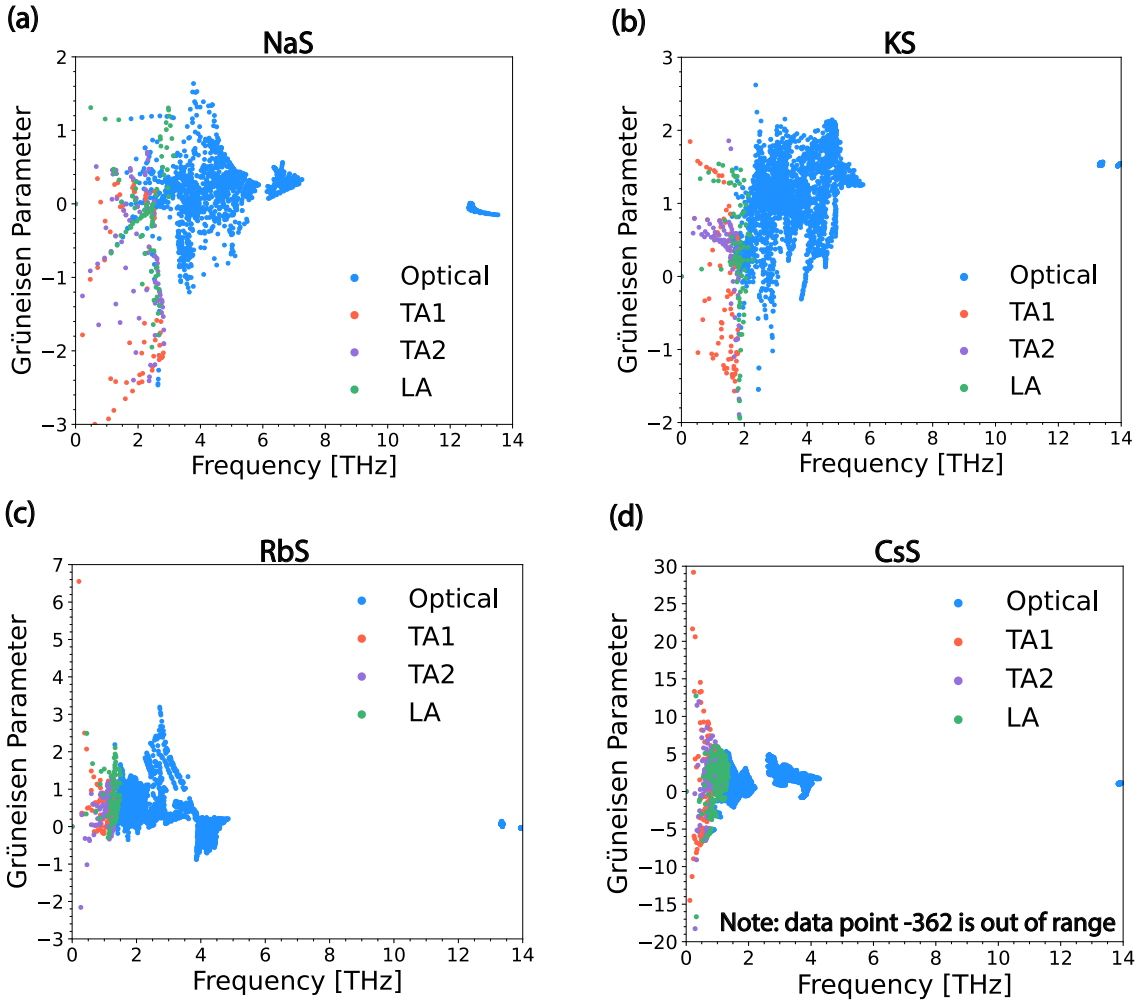

Figure 2.12: **Grüneisen Parameter for XS Materials.** (a) Grüneisen parameters for NaS. (b) Grüneisen parameters for KS. (c) Grüneisen parameters for RbS. (d) Grüneisen parameters for CsS.

2.5.5 Electronic Properties of the XS Family

Table 5: Effective mass of the XS (X=Na, K, Rb, Cs) family. The definition of the **k**-points are based on the reciprocal lattice of the primitive cell.

Material	Carrier	Direction	Effective Mass (m_0)	Comment
NaS	electron	$\Gamma(0,0,0) \rightarrow K(\frac{1}{3},\frac{1}{3},0)$	0.55	In-plane
	electron	$\Gamma(0,0,0) \rightarrow A(0,0,0.5)$	0.39	Out-of-plane
	hole	$\Gamma(0,0,0) \rightarrow K(\frac{1}{3},\frac{1}{3},0)$	13.88	
	hole	$\Gamma(0,0,0) \rightarrow M(0.5,0,0)$	15.92	
	hole	$\Gamma(0,0,0) \rightarrow A(0,0,0.5)$	1.43	Out-of-plane
KS	electron	$A(0,0,0.5) \rightarrow R(0.5,0.5,0.5)$	1.48	In-plane
	electron	$A(0,0,0.5) \rightarrow \Gamma(0,0,0)$	0.24	Out-of-plane
	hole	$\Gamma(0,0,0) \rightarrow K(0.5,0.5,0)$	3.41	In-plane
	hole	$\Gamma(0,0,0) \rightarrow A(0,0,0.5)$	1.21	Out-of-plane
RbS	electron	$A(0,0,0.5) \rightarrow R(0.5,0.5,0.5)$	1.18	In-plane
	electron	$A(0,0,0.5) \rightarrow \Gamma(0,0,0)$	0.29	Out-of-plane
	hole	$\Gamma(0,0,0) \rightarrow M(0.5,0,0)$	6.82	
	hole	$\Gamma(0,0,0) \rightarrow K(0.5,0.5,0)$	6.87	
	hole	$\Gamma(0,0,0) \rightarrow A(0,0,0.5)$	1.46	Out-of-plane
CsS	electron	$C(0.22,-0.22,0.22) \rightarrow W(\frac{1}{4},\frac{1}{4},\frac{1}{4})$	0.82	
	electron	$C(0.22,-0.22,0.22) \rightarrow T(0,0,0.5)$	0.86	
	electron	$C(0.22,-0.22,0.22) \rightarrow R(0,0.5,0)$	0.71	
	hole	$\Gamma(0,0,0) \rightarrow X(-0.5,0.5,0.5)$	2.27	
	hole	$\Gamma(0,0,0) \rightarrow Y(0.5,-0.5,0.5)$	1.68	
	hole	$\Gamma(0,0,0) \rightarrow Z(0.5,0.5,-0.5)$	1.92	

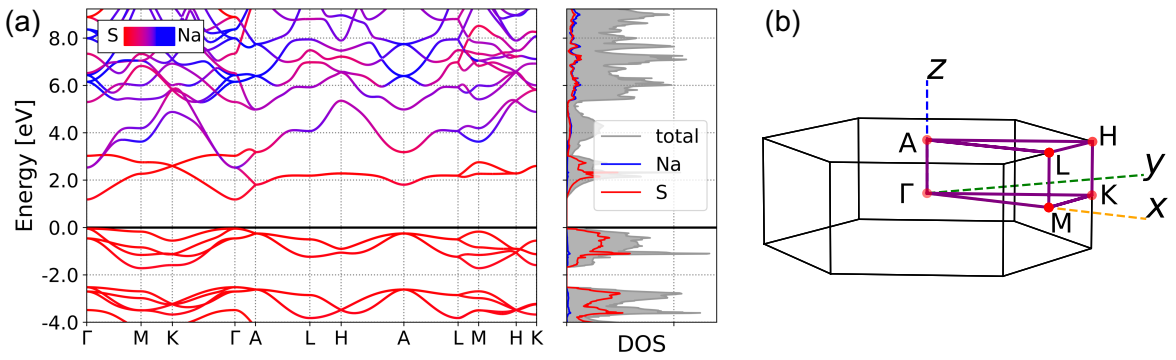

Figure 2.13: (a) Band structure and density of states for NaS. (b) Special **k**-points in the first Brillouin Zone in NaS.

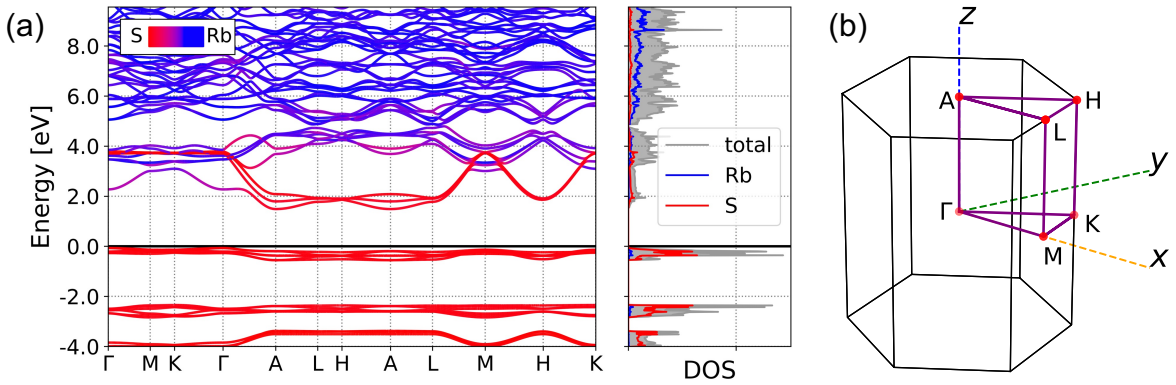

Figure 2.14: (a) Band structure and density of states for RbS. (b) Special **k**-points in the first Brillouin Zone in RbS.

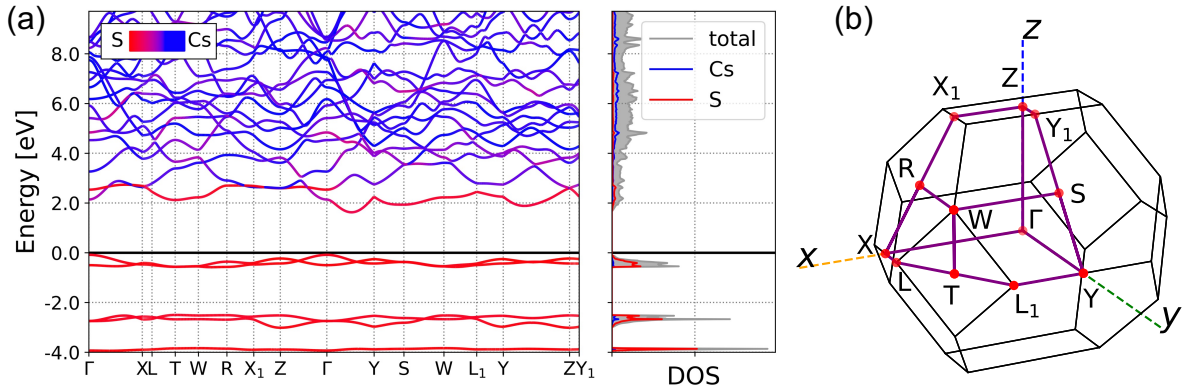

Figure 2.15: (a) Band structure and density of states for CsS. (b) Special **k**-points in the first Brillouin Zone in CsS.

Supporting Information

A screened list of binary semiconductors with ABVBs can be found here `https://pubs.acs.org/doi/suppl/10.1021/jacs.3c05091/suppl_file/ja3c05091_si_001.pdf`.

Permissions and Attributions

The content of chapter 2 is the result of a collaboration with Yubi Chen and Bolin Liao, and has previously appeared in the *"Lattice Dynamics and Thermal Transport in Semiconductors with Anti-Bonding Valence Bands"*, Journal of American Chemical Society, 145 (33), 18506-18515 (2023). It is reproduced here with the permission of American Chemical Society.

Chapter 3

Enhancing Magnetocaloric Material Discovery using Machine Learning and Large Language Models

3.1 Introduction

Solid-state refrigeration presents a groundbreaking avenue driven by the pressing need for more sustainable and efficient cooling technologies [25, 134, 135, 136]. Specifically, magnetic refrigeration shows promise in solid-state cooling applications across a wide temperature range. Magnetic refrigeration is based on the magnetocaloric effect (MCE), where the entropy of a magnetic material (the "magnetic refrigerant") is controlled by an external magnetic field [25]. When an external magnetic field is applied, the microscopic magnetic moments are aligned with the field. As a result, the entropy associated with the magnetic moments in the material is reduced. In the process, either heat is released under an isothermal condition or the temperature of the magnetic refrigerant is increased under an adiabatic condition. When the magnetic field is removed, the effect is reversed.

These processes can be combined to form a thermodynamic cooling cycle, whose cooling capacity is determined by the isothermal magnetocaloric entropy change (ΔS_M) and the adiabatic temperature change (ΔT_{ad}) during the magnetization and demagnetization processes. This cycle resembles a Carnot cycle on the S-T diagram, promising close-to-ideal efficiency [137]. Compared with other cooling technologies, magnetic refrigeration has received renewed interest thanks to its solid-state nature, cryogen-free operation, reliability, compactness, and capability to reach close-to-Carnot efficiency at cryogenic temperatures [138]. For these reasons, magnetic refrigeration has been widely applied for applications in space missions [139], scientific instrumentation [140], and as an environmentally friendly alternative to vapor compression-based cooling [141]. However, most previous investigations of magnetic refrigeration have focused on either cryogenic temperatures below a few K or near room temperature, while many technologically important applications require efficient cooling in the temperature range between 10 K and 100 K, such as hydrogen and methane liquefaction [142] and cryogenic carbon capture [143].

To achieve efficient magnetic refrigeration in this temperature range, one key requirement is to identify new high-performance magnetic refrigerants with their Curie temperature (T_C) in this range, since both ΔS_M and ΔT_{ad} are typically maximized near T_C. Therefore, T_C signifies the ideal working temperature for an MCE material and stands as a crucial parameter for identifying new magnetic refrigerants. In addition, large magnetic moments and favorable exchange interactions [144] are fundamental prerequisites for a material to exhibit favorable MCE characteristics. In this context, rare-earth-based alloys are state-of-the-art candidates for magnetocaloric refrigeration in the intermediate temperature range. For example, heavy rare-earth-based MCE materials, such as HoB_2, [29, 52] $ErCo_2$, [53] $DyAl_2$, [54, 55] and $ErAl_2$, [54] exhibit significant ΔS_M and ΔT_{ad} values in this temperature range due to their large magnetic moments. Consequently, they are actively researched for applications like hydrogen liquefaction

because they have T_C within the required range [56, 57, 58]. Their typical isothermal entropy change ΔS_M is around 20-30 $\mathrm{J\,kg^{-1}K^{-1}}$ and adiabatic temperature change ΔT_{ad} is around 10 K for an applied magnetic field of 5 T. Historically, these MCE materials were discovered mainly through an experimental synthesis and characterization process, which can be expensive, time-consuming, and often relies on trial and error.

To make further progress, more efficient methods to discover and verify new high-performance MCE materials are needed for refrigeration applications in various temperature ranges. Along this line, high-throughput computational material searches were previously conducted [64, 60]. For example, Bocarsly et al. used the strength of spin-lattice coupling in a material, which can be effectively evaluated at the ground-state density functional theory (DFT) level, as a proxy to search for new magnetocaloric materials [64]. In addition, previous studies have used various machine learning-based methods to predict suitable candidates for magnetic refrigeration from the aspects of compositional and structural changes [61] and optimization to enhance key performance metrics while reducing material cost [62, 63]. Suitable candidates such as HoB_2 and Co-doped Fe_2P-type compounds were identified for this purpose by data mining and machine learning [29, 66]. However, these pioneering studies have been limited in the size of accessible MCE datasets and have lacked an efficient way to verify the MCE performance of predicted materials to close the loop for material discovery. As a result, new MCE materials of practical value that have been identified through these approaches have been limited.

In the realm of materials science, there is a vast and expanding collection of unstructured data in the form of journal articles, patents, and theses. This accumulation and rapid publication pace overwhelm researchers' ability to fully utilize the available information. Converting these data into structured materials databases would facilitate data-driven research aimed at discovering new materials, yet manual compilation is not only labor-intensive, but also susceptible to errors. Fortunately, advancements in natural

language processing (NLP) have enabled the automated extraction of data from extensive document collections, paving the way for creating detailed materials property databases and discovering new materials [145, 146]. For example, in a previous study, Court *et al.* have established an auto-generated database of 39,822 records containing magnetic materials and their associated Curie or Néel magnetic phase transition temperatures using NLP to extract relevant information from the literature [147]. In a subsequent study, they used an NLP toolkit to generate a material database from published literature containing 2,910 entries with their MCE properties and then used the data to train a deep neural network to predict MCE properties of new materials [65]. They identified six new promising MCE materials with this innovative approach. However, the validation of their results still relied on experiments that can be the bottleneck of the pipeline.

Two recent advancements have provided new opportunities to further improve the pipeline proposed by Court *et al.* [65] to efficiently and accurately identify new MCE materials. One is the rapid rise of large language models (LLMs), particularly those in the Generative Pre-trained Transformer (GPT) series [148], which have demonstrated exceptional capabilities in generating coherent and contextually appropriate text and have also introduced advanced features such as summarizing articles, generating informative charts, and mapping relationships between words. The emergence of LLMs enables efficient extraction of material properties from a large amount of published literature [32, 149]. The other development is computational methods based on atomistic spin dynamics (ASD) simulations with exchange interaction coefficients computed from first principles [150, 151, 144]. These methods allow an accurate evaluation of MCE properties based solely on the crystal structure and chemical composition of materials and can close the loop in high-throughput and machine learning-based computational material searches.

In this work, we exploit GPT's remarkable abilities to summarize texts to complete

an auto-generated MCE material database with 6615 entries based on openly accessible abstracts of published literature. After careful curation and validation, we used high-quality data from the database to train a neural network that can predict the isothermal entropy change of ferromagnetic materials with high fidelity only based on their chemical composition. With this neural network model, we efficiently screened binary and ternary magnetic compounds in the large online database, the Materials Project [108], and identified a list of promising MCE materials that have not been studied before. Finally, we used DFT-based first-principles simulation and ASD to evaluate the MCE properties of the top-ranked candidates predicted by the neural network and identified 11 new promising MCE materials for the less explored intermediate temperature range (10 K to 100 K). Importantly, the calculated results can be added back to the MCE database which can further improve the performance of the neural network in the fashion of a closed feedback loop. A schematic of our workflow is shown in Fig. 3.1 with more detailed information about each key step summarized in Fig. 3.2. Our work demonstrates the vast potential of using LLMs to extract structured information from published literature and utilize the extracted data to drive new discoveries of functional materials.

3.2 Methods

3.2.1 Database Creation using ChatGPT

To further the discovery of new MCE materials through machine learning, it is crucial to establish a comprehensive and high-quality database containing the MCE properties of materials that can be then used as the training dataset for a machine learning model. Fortunately, the rapid advancement of LLMs, particularly with the introduction of ChatGPT by OpenAI in 2022, offers a promising solution to automatically generate a database

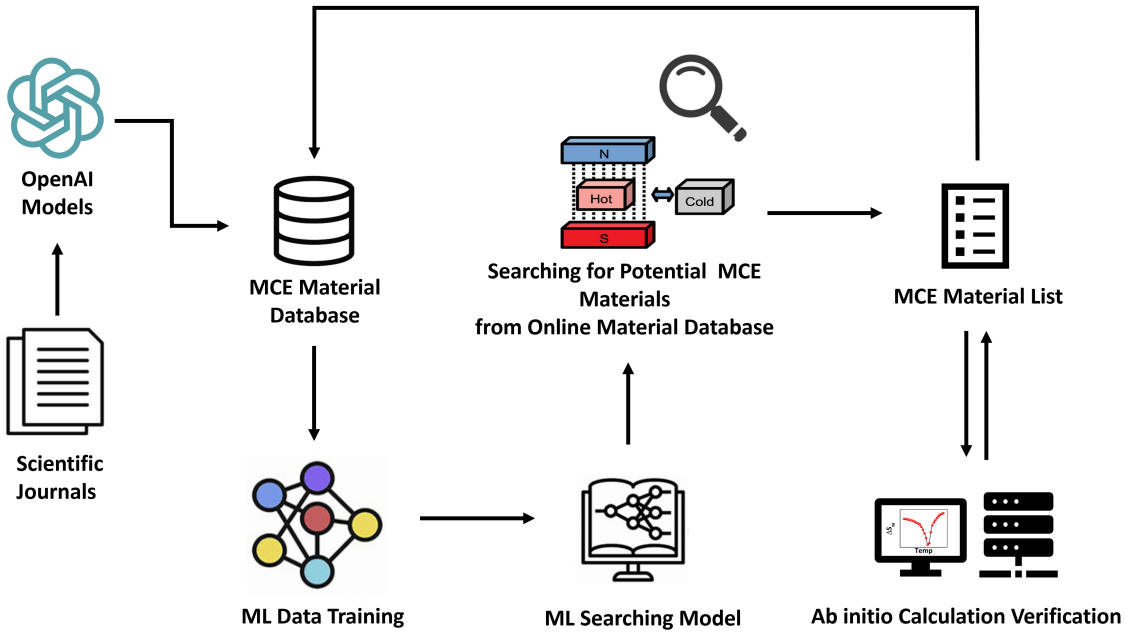

Figure 3.1: **Flowchart of the magnetocaloric material discovery pipeline**. Existing literature is analyzed by a large language model to generate a database containing the MCE properties of a large number of materials. The database is then curated and refined to serve as the training dataset for a machine learning model. The machine learning model is then used to screen materials in the Materials Project database to identify new promising MCE materials, whose properties are verified by ab initio DFT and ASD simulations. The simulation results are added back to the material database to form a closed loop to refine the process.

based on the large body of published literature [152]. The process for creating our MCE database using ChatGPT is detailed below:

Data Collection: We began by selecting the first 8,000 most relevant abstracts from the Web of Science from more than 350 journals, ranking all scientific papers by their relevance to the MCE properties of materials. The time span of the publication date extends from 1993 to 2023. These abstracts were then converted into CSV files containing both the text of the abstract and publication details. To streamline this process, we focused on papers that provided useful data, such as those containing one or more chemical compounds and their corresponding MCE properties expressed in numerical values.

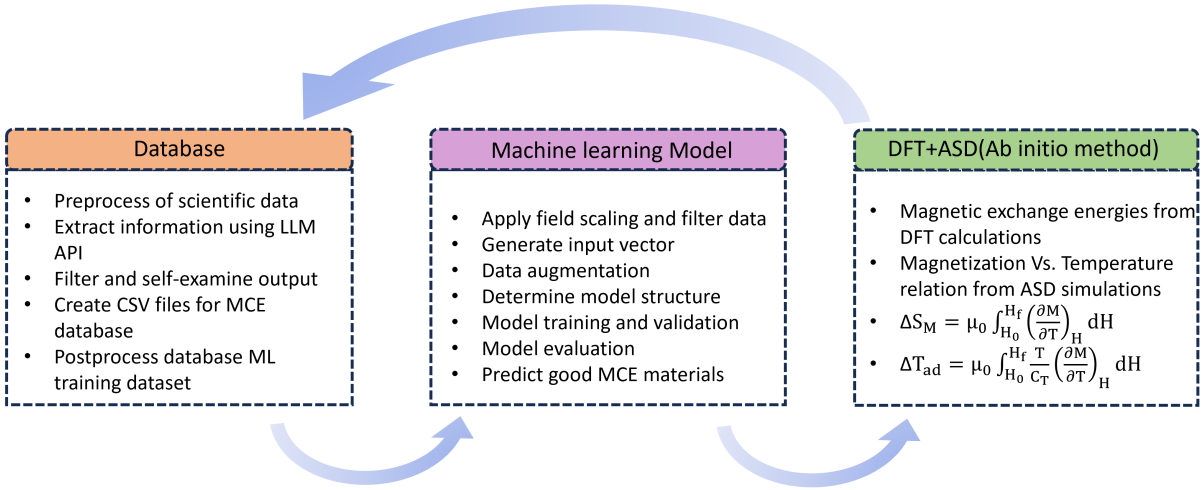

Figure 3.2: **Detailed information on the key steps of the magnetocaloric material discovery pipeline.**

To reduce the number of records to a more tractable quantity for further analysis, we developed a filter pipeline that employs prompts to determine the potential usefulness of data within an abstract. This filtering process is both accurate and efficient, as commercialized LLMs like GPT can readily identify the presence of chemical compounds or numerical values in the text. Abstracts that only offer general descriptions or values without specific materials were excluded.

Prompt Engineering and Self-Examination: Utilizing the OpenAI's LLM API, we transformed abstracts into tables, highlighting material compositions and their properties. This involved using a multi-shot prompt strategy [153, 154], which increases the accuracy of data extraction by providing the LLM with multiple examples. The tables generated were generally accurate. GPT models, like other LLM models, exhibit a phenomenon called hallucination, wherein they can generate low-likelihood outputs that deviate from the expected output based on the input context and the true data distribution [155]. To address this potential issue, we applied a self-examination pipeline that aims to reduce the likelihood of errors due to hallucinations, such as the inclusion of unmentioned materials, overlooking properties, or confusing property names. The self-

examination process involves an additional prompt to verify the consistency between the generated table and the original abstract, specifically to counteract the LLM's tendency to make educated guesses by removing numerical values not provided in the abstract. These steps were conducted separately, guiding the LLM to focus on task reasoning rather than relying on its vast knowledge base.

Filtering after Natural Language Processing: After extracting relevant material formulations and MCE properties from the literature using GPT, we established a database with MCE properties for 6615 materials. Starting from this collection, we refined the dataset by filtering out organic materials and alloyed materials to limit our material discovery process to intrinsic inorganic materials as the starting step, which leads to a reduced database with 1426 entries. In principle, however, the material search can be expanded later to include more complex materials in our future work.

Database Creation and Mapping to the Materials Project: Our objective was to create a one-to-one match between the materials listed in our filtered database and those cataloged in the esteemed Materials Project (MP) database [108]. This approach was chosen because the MP database provides a wealth of physical properties data (notably, without the MCE properties), which, when combined with our MCE data, facilitates their use in machine learning for materials discovery. Ultimately, we successfully matched 752 materials in our database with their corresponding Material Project IDs. To enhance accessibility and utility, we developed an HTML version of the database [1]. This allows users to efficiently search for reported MCE materials using various criteria, such as chemical elements, or to organize materials by their name, MCE properties, or Curie temperatures.

Evaluation: We manually reviewed 200 randomly chosen samples from each test set to assess the database's accuracy. The database was estimated to have an accuracy of

[1] `https://github.com/RunqingYang1996/MCEdatabase`

around 90%. A record is considered valid only if the name of one material matches the correct magnetocaloric entropy change (ΔS_M) and T_C values cited in the input scientific text. Information about our database is summarized in Table 1.

Table 1: Information about the created and filtered MCE databases.

Number of Records	Database Information	Estimated Precision
6615	Material name, ΔS_M and/or T_C. Organic and alloyed materials are included.	85%
1426	Material name, ΔS_M and/or T_C. Only intrinsic inorganic materials are included.	90%
752	Material name, ΔS_M and T_C, and Materials Project ID.	90%
752	Human-corrected version of the 752-record database for machine learning training	99%

3.2.2 Machine Learning Model Construction and Training

Our next step is to train a machining learning model using the database constructed by the LLM to predict MCE properties of materials based on their chemical composition. One challenge arises that the reported ΔS_M values were measured across varying magnetic field strengths in different papers, rendering them incomparable. To address this, we implemented a linear scaling based on a commonly used field strength of 5 T to approximate the field dependence of ΔS_M as established in measurements [156]. Another issue is that the dataset contains a notable number of duplicates due to measurements by different groups or at different field strengths. We kept the maximum value among the duplicates after scaling to make ΔS_M values consistent for training. To leverage data efficiency and model complexity, we limited the training data to include only materials with less than or equal to 8 atoms in the unit cell, resulting in 229 unique materials in the training set. We adopted a feedforward neural network with 3 hidden layers, each with 16 neurons. The optimal network structure was determined via a grid search of different numbers of layers and neurons per layer. Details of the process can be found in Supplementary Information (SI) Sec. 3.5. The input vector contains 4 features with

a total vector size of 11: 8 integers representing the atoms in the unit cell, and 3 floats for volume, density, and total magnetization, respectively, which can be extracted from the Materials Project database. In general, neural networks require larger amounts of high-quality training data, enabling them to learn the underlying patterns and relationships. In our case, a major challenge is that we achieved high training data quality at the expense of the total number of records. In this study, we tackled this by employing data augmentation. Since the order of atomic numbers in the vector representation can be arbitrary, permutations of atomic numbers essentially represent the same material. Hence, we shuffled the elements of the atomic number vector to create multiple representations and augment the size of the training set.

In the model, we generated 100 "representations" of each unique material as training data. Test runs that validated the proof of concept are included in SI 3.5. The training process aims to minimize the mean squared error (MSE), which is more susceptible to outliers because it penalizes larger errors more significantly. In this context, MSE is suitable for the purpose of identifying candidates with large ΔS_M values. Standardization is applied to the dataset, which is a common preprocessing step in machine learning to help with algorithm convergence. It ensures that each column of input features and output has zero mean and unit variance. The target of the model, standardized ΔS_M values, is directly used to rank the candidate materials. The predictions can be converted back to the original scale by an inverse transformation. Fig. 3.3 presents the training curves and an evaluation of the adopted model. The first two panels show the loss function (MSE) and the mean absolute error (MAE) of the training process, indicating the model's accuracy and convergence. Notably, the consistently low values of both MSE and MAE across the panels affirm the robustness and accuracy of the adopted model. The third panel illustrates the correlation between predicted ΔS_M and true ΔS_M values of the dataset, showing the model's predictive capabilities. For each data point, the atomic

number vector of the input representation was shuffled 10 times, and the average of the 10 predictions was plotted. The high correlation coefficient of 0.99 underscores a strong alignment between the predicted values by the neural network and the ground truth. This impressive correlation emphasizes the model's reliability in capturing patterns within the data and reinforces its suitability for accurate predictions.

We executed the model 10 times to assess the consistency of rankings and produce the final ranking. The ranking from each run is recorded in a sorted list for each material, and this sorted list is used for the final ranking based on its length and each ranking. We evaluated the model performance by the "recall" of the ranked list [157]. Recall is a performance metric in machine learning that measures the true positive rate of a model's prediction. In other words, it measures the proportion of actual positive cases that the model correctly identifies as positive. For example, "recall@10" measures that, among the top 10 candidates predicted by the model, the percentage that accurately aligns with the top 10 candidates according to the ground truth. Table 2 shows the recall when considering the top K candidates. "Max ranking" indicates the number of top candidates considered in producing the sorted list, which can only be smaller or equal to K. By examining the recalls at various K values, it becomes evident that our model demonstrates consistent performance across different ranking strategies.

Table 2: Recall of top K materials ranked by ΔS_M value.

Max Ranking / K	10	25	50
10	1.0	1.0	1.0
25	-	0.96	0.96
50	-	-	0.96

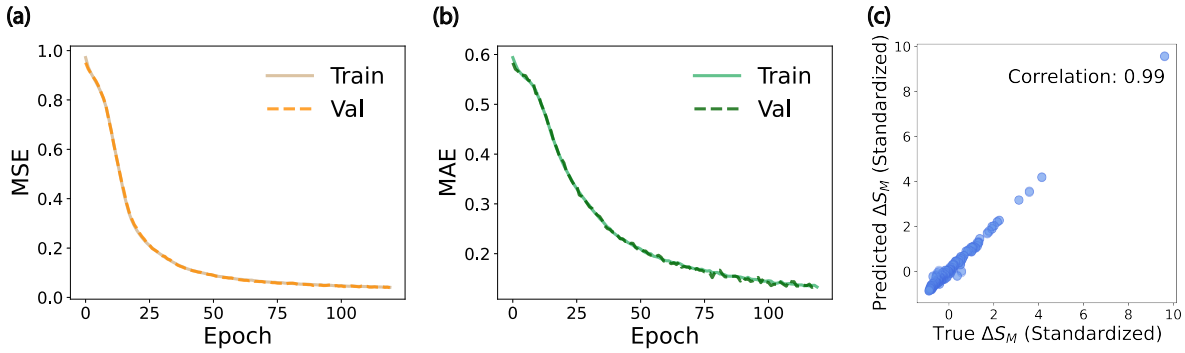

Figure 3.3: **Training and the performance metric of the machine learning model.** Training curves of (a) the mean-squared error (MSE) and (b) the mean-absolute error (MAE) of the machine learning model. Here the metrics for the training set and the validation set are plotted in solid and dashed lines, respectively. (c) Correlation plot of the predicted standardized ΔS_M by the trained machine learning model versus the true standardized ΔS_M in the database.

3.2.3 DFT and ASD Simulations

Once new promising MCE materials are predicted by the machine learning model, we use DFT and ASD simulations to directly compute their MCE properties (ΔS_M and ΔT_{ad}) and verify the performance of the machine learning model. For each magnetic material, the magnetic exchange energies (J_{ij}) were estimated with DFT calculations using the Vienna ab initio simulation package (VASP) based on the projected augmented wave pseudopotentials [158]. The Perdew-Burke-Ernzerhof form of generalized gradient approximation (PBE-GGA) [111] was used for structural optimization. All of our calculations utilized a plane-wave cut-off energy of 600 eV. The energy and force convergence criteria were set to be 1×10^{-5} and 0.01 eV/Å, respectively. Monkhorst-Pack [159] **k**-point grid of ($20 \times 20 \times 20$) was employed for cubic GdMg, and similar density of **k**-point grids were utilized to sample the Brillouin zone for all other compounds. The robust correlation among the localized 'f' electrons was systematically addressed via GGA+U calculations. Specifically, the parameters U and J were set to 7 eV and 1 eV, respectively, for all rare-earth elements in our calculations, as these values are commonly

used in the literature [160, 161]. Three nearest-neighbor J_{ij} values were estimated in each case and subsequently utilized for ASD simulations in further steps.

The parameterized atomistic spin model was employed for ASD simulations using VAMPIRE simulation software [162]. To determine the T_C, the Monte Carlo method was utilized. This involved executing 10,000 steps at each temperature, allowing the system of $30 \times 30 \times 30$ unit cells with in-plane periodic boundary conditions to reach equilibrium. Subsequently, a statistical average was obtained over 10,000 steps to extract the mean magnetization. Temperature-dependent magnetization curves were calculated employing the spin dynamics approach and the Heun integration scheme [163]. Additionally, the demagnetization field induced by the atomistic spins themselves was accounted for. The external magnetic field was systematically increased to 5 T with an incremental step of 1 T. The ASD simulation was conducted to generate the magnetization versus temperature (M vs. T) curves. Figure 3.4 represents the M vs. T curves for three representative materials i.e. EuB_6, $EuLiH_3$, and GdMg. From these temperature-dependent magnetization curves, T_C values were estimated by fitting the curves with $M(T) = \left(1 - \frac{T}{T_C}\right)^{\beta}$, where β is the critical exponent. The M vs. T curves for fields ranging from 0 to 5 T were calculated with steps of 1 T to evaluate ΔS_M using the Maxwell relation [25]:

$$\Delta S_M = \mu_0 \int_{H_i}^{H_f} \left(\frac{\partial M}{\partial T}\right)_H dH, \tag{3.1}$$

where H_i and H_f represent the initial and final magnetic fields, respectively, and μ_0 denotes the Bohr magneton. For further evaluation of the MCE performance of the predicted compounds, the total specific heat (C_T) was estimated as a combination of phononic (C_p), magnonic (C_m), and electronic (C_e) contributions, i.e.,

$$C_T = C_m + C_p + C_e. \tag{3.2}$$

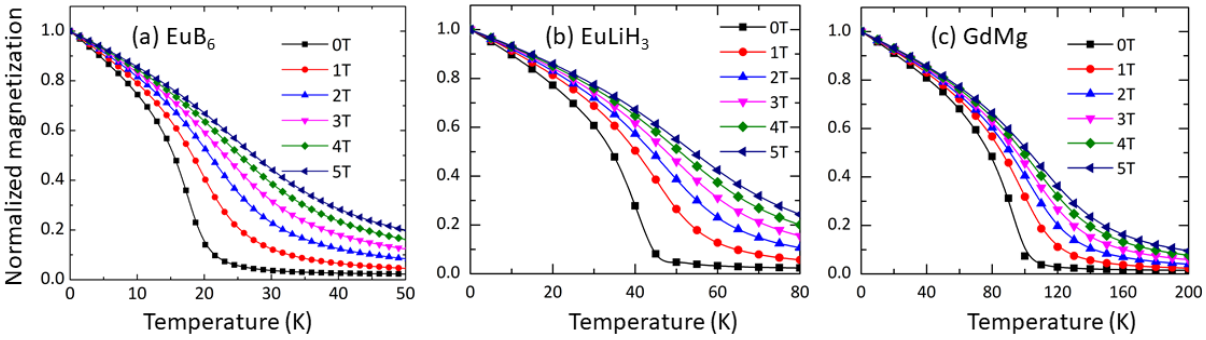

Figure 3.4: **Simulated temperature and field-dependent magnetization in three representative compounds predicted by the machine learning model.**

First, we begin with the magnetic energy (U) of the magnetic compound to calculate the C_m i.e.

$$U = -\sum_{ij} \lambda_{ij} \vec{J}_i \cdot \vec{J}_j - \sum_i g\mu_B \vec{J}_i \cdot h^{ext}, \tag{3.3}$$

where the first term represents the spin-spin exchange interaction, where λ_{ij} is the exchange interaction parameter and J_i is the total angular momentum of the magnetic ions. The second term represents the Zeeman interaction of the total angular momentum with an external magnetic field (h^{ext}). For a given temperature, the mean value of the energy $\langle U \rangle$ is calculated by:

$$\langle U \rangle = \frac{1}{(N_C - N_0)} \sum_{i>N_0}^{N_C} U_i, \tag{3.4}$$

where N_C represents the total number of Monte Carlo cycles and N_0 is the number of Monte Carlo cycles used for thermalization. The mean square energy $\langle U^2 \rangle$ is obtained by a relation similar to Equation 3.4. Now, the magnetic part of the heat capacity can be calculated as

$$C_m(T, h^{ext}) = \frac{\langle U^2 \rangle - \langle U \rangle^2}{k_B T^2}, \tag{3.5}$$

where k_B is the Boltzmann constant. In this study, the Monte-Carlo simulations were performed using the VAMPIRE simulation software [162], which produces the magnetic heat capacity as one of the outputs.

Following the determination of phonon frequencies through the frozen phonon method, constant volume phonon heat capacity (C_V) can be calculated using the following equation as implemented in the Phonopy code [164, 165]:

$$C_V = \sum_{qj} C_{qj} = \sum_{qj} k_B \left(\frac{\hbar\omega_{qj}}{k_B T}\right)^2 \frac{\exp(\hbar\omega_{qj}/k_B T)}{[\exp(\hbar\omega_{qj}/k_B T) - 1]^2}, \tag{3.6}$$

where q is the wave vector, j is the band index, ω_{qj} is the phonon frequency, and $\hbar$ is the reduced Planck's constant. We used the calculated C_V as an approximation for the constant-pressure heat capacity C_p, since in the low-temperature regime of solids, $C_V \approx C_p$ [166].

Lastly, C_e is given by:

$$C_e = \frac{\pi^2 k_B^2 D(E_F)}{3} T = \gamma_e T, \tag{3.7}$$

where $\frac{\pi^2 k_B^2 D(E_F)}{3}$ represents the electronic heat-capacity coefficient, and $D(E_F)$ represents the electronic density of states at the Fermi energy E_F. Next, the ΔT_{ad} was estimated using the calculated C_T by employing the following equation:

$$\Delta T_{ad} = \mu_0 \int_{H_i}^{H_f} \frac{T}{C_T} \left(\frac{\partial M}{\partial T}\right)_H dH. \tag{3.8}$$

3.3 Results and Discussion

3.3.1 Database Generation and Machine Learning Model

We successfully created a database with high accuracy using the LLM API, demonstrating the efficacy of our pipeline. The potential for hallucination (i.e., incorrect ΔS_M or T_C values) was avoided through the self-examination process. The final database, containing 752 entries, is machine learning-friendly due to the clarity of material names

and the ease of finding additional physical properties by cross-referencing the Materials Project database. However, there are still key issues that limit the precision of the database. Firstly, the text format cannot always accurately identify mathematical or physical symbols, leading to the loss of key information (e.g. numerical value or unit). Secondly, the LLM API sometimes fails to make correct judgments when the text mentions both Curie temperature and Néel temperature. As a result, some of the Néel temperature records within the database have been incorrectly identified as Curie temperatures, which is an intrinsic issue with the LLM. Compared to this, the LLM has high accuracy for ΔS_M values for each record. We used a human-corrected version of the database for higher accuracy training of the machine learning model.

One strength of our model is that we added additional physical information to our input vectors obtained from the Materials Project database during the training process. Physical information is added by containing total magnetization, volume, and density information from the Materials Project and the spatial partition of the element distribution. The results showed robust performance-based MSE and MAE. We also did a recall evaluation of the model. In industrial applications, recall evaluation is frequently used where the emphasis is on identifying and capturing as many true positive cases as possible, even at the expense of some false positives. By examining the recalls at various K values, it becomes evident that our model can consistently rank materials based on their ΔS_M values. Additionally, our model stands out for its remarkable efficiency. Unlike many other studies that rely on complex model architectures or require extensive datasets with hundreds of features, this model achieves impressive results with a surprisingly simple structure. It utilizes a neural network with only 3 layers and 16 neurons and leverages a basic input vector of just 4 essential features with a total length of 11. This combination of simplicity and effectiveness makes this model a compelling choice for faster iteration and prediction.

The frequency distribution of ΔS_M and T_C values of the training dataset are shown in SI Fig. S1 3.5. It is noteworthy that 70.0% of materials in the training dataset exhibit a T_C below 100 K. This prevalent feature is crucial for the model's capacity to identify additional candidates with a T_C within the desired range (10 K to 100 K). The frequency distribution of rare-earth elements in our training dataset in SI Figure c 3.7 shows that 81.6% of reported MCE candidates in the dataset contain heavy rare-earth elements. This prior knowledge and observation motivated us to generate a screening list from the Materials Project by requiring the presence of heavy rare-earth elements: Gd, Tb, Dy, Ho, Er, Tm, and Eu. Using this criterion, we found 1124 stable materials containing 2 to 8 atoms in the unit cell. We separated them into three lists based on the number of elements: 370 candidates contain two elements, 711 candidates contain three elements, and 43 candidates contain at least four elements. The machine learning model provides one ranked list for each group. We picked the top 50 materials predicted by the machine learning model for further studies using DFT and ASD.

3.3.2 DFT and ASD Calculations

The primary objective of machine learning is material prediction; therefore, it is necessary to assess the quality of the predictive model. Hence, we validated the MCE response of the predicted compounds through DFT and ASD simulations. Initially, the predicted structures were optimized within the spin-polarized PBE-DFT level of theory. The total energies of magnetic structures with various combinations of in-plane and out-of-plane magnetic orderings were compared to determine the magnetic ground state. Materials exhibiting ferromagnetic ground states were further analyzed for their MCE properties, as they typically demonstrate direct MCE with $\Delta S_M < 0$ and $\Delta T_{ad} > 0$, [167, 168] indicating that the sample heats up when an external magnetic field is applied

adiabatically. Materials with a direct MCE effect typically exhibit higher magnitudes of magnetic entropy and adiabatic temperature changes compared to the inverse MCE observed in antiferromagnetic or ferrimagnetic materials [169].

After running a magnetic ground state simulation based on DFT of the top 50 candidates predicted by our machine learning model, we identified 11 ferromagnetic compounds with their T_C in the range 10 K to 100 K that have not been studied before for their MCE properties. We then conducted detailed DFT and ASD simulation of their MCE properties. The simulated temperature and field-dependent magnetization and the ΔS_M values of a few representative compounds are shown in Fig. 3.4 and Fig. 3.5. The crystal structure symmetry, T_C, ΔS_M, and ΔT_{ad} values evaluated using DFT and ASD for the machine learning-predicted compounds, are provided in Table 3. The estimated T_C values closely match the reported measured values, considering that our simulations do not account for other effects such as anisotropy energy and crystal effective field effects.

Our computational predictions reveal a diverse array of magnetocaloric materials characterized by their T_C, ranging from the cryogenic temperatures of liquid helium (4.2 K) to near the boiling point of liquid nitrogen (77 K) (Table 3). These materials can be systematically classified into three distinct subgroups: lower (10 - 30 K), medium (30 - 50 K), and higher (50 - 90 K) T_C materials. Notably, the materials in the lower T_C subgroup demonstrate exceptional potential for applications such as hydrogen liquefaction (typically around 20 K) and low-temperature magnetic cooling. A notable candidate among these lower T_C materials is EuB_6, which exhibits a theoretically predicted ΔS_M of approximately 48 $\mathrm{J\,kg^{-1}K^{-1}}$, surpassing, to the best of our current knowledge, all other known materials in terms of mass-based performance within this temperature range [29]. However, when considering the isothermal entropy change per volume, the ΔS_M of EuB_6 reduces to approximately 0.24 $\mathrm{J\,cm^{-3}K^{-1}}$, a value comparatively lower than that of HoB_2, which exhibits a ΔS_M of approximately 0.36 $\mathrm{J\,cm^{-3}K^{-1}}$ [29]. This disparity arises pri-

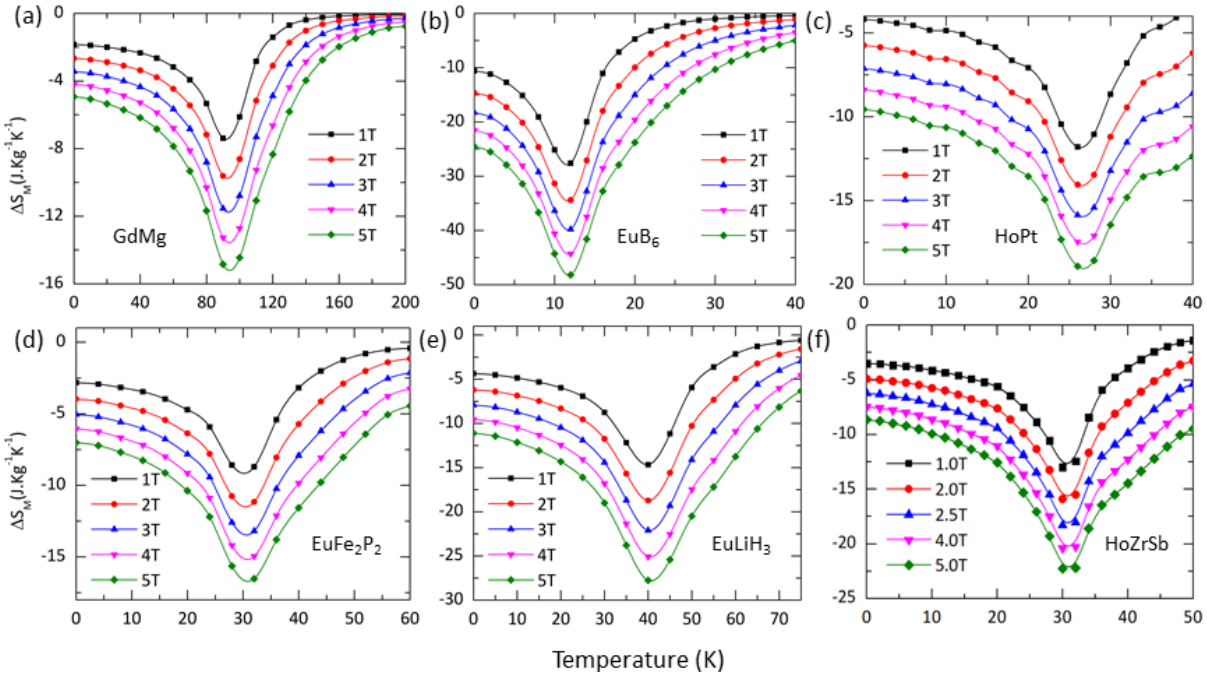

Figure 3.5: **Simulated temperature and field-dependent magnetocaloric entropy change in six representative compounds predicted by the machine learning model.**

marily due to the lower density of EuB_6 (5.01 g cm^{-3}) compared to HoB_2 (8.68 g cm^{-3}). Additionally, the estimated ΔT_{ad} for EuB_6 is approximately 8.1 K, a figure comparable to other materials with similar transition temperatures [29]. In the lower T_C subgroup, HoPt exhibits the highest ΔS_M value in volumetric units, reaching 0.25 J cm^{-3}K^{-1}.

In the medium-T_C range, $EuLiH_3$ exhibits a notable ΔS_M value of 28 J kg^{-1}K^{-1} (0.15 J cm^{-3}K^{-1}) and a ΔT_{ad} of 6.7 K at approximately 40 K. These values are comparable to the values observed for TbN at 44 K ($\Delta S_M \approx 21$ J kg^{-1}K^{-1} ≈ 0.20 J cm^{-3}K^{-1}, $\Delta T_{ad} \approx$ 8 K in the $\langle 001 \rangle$ direction) [177, 178]. Moving to the higher-T_C range, DyZrSb displays the largest ΔS_M at 52 K, while GdMg exhibits the smallest at 90 K (refer to Table 3). Additionally, it is noteworthy that the predicted materials include a series with the general formula RZrSb (where R = Gd, Tb, Dy, and Ho), spanning a wide temperature range from 28 K to 70 K. Among these, HoZrSb stands out with the highest estimated

Table 3: The crystal structure, estimated (reported) T_C, and MCE properties for a field change of 5 T of compounds predicted by the machine learning model.

Compound	Structure	T_C (K)	ΔS_M ($\mathrm{J\,kg^{-1}K^{-1}}$)	ΔS_M ($\mathrm{J\,cm^{-3}K^{-1}}$)	ΔT_{ad} (K)
EuB_6	Cubic ($Pm\bar{3}m$)	12 (16)[a]	48	0.24	8.1
TbCoC	Tetragonal (P4/mmm)	16 (22)[b]	20	0.17	5.9
ErPt	Orthorhombic (Pnma)	24 (16)[c]	14	0.20	4.5
HoPt	Orthorhombic (Pnma)	26 (16)[c]	18	0.25	5.3
HoZrSb	Tetragonal (I4/mmm)	28 (22)[d]	21	0.18	6.2
$EuFe_2P_2$	Tetragonal (I4/mmm)	32 (30)[e]	17	0.12	7.2
$EuLiH_3$	Cubic ($Pm\bar{3}m$)	40 (38)[f]	28	0.15	6.7
GdZrSb	Tetragonal (I4/mmm)	45 (37)[d]	14	0.11	4.6
DyZrSb	Tetragonal (I4/mmm)	52 (45)[d]	19	0.16	5.4
TbZrSb	Tetragonal (I4/mmm)	70 (61)[d]	16	0.13	5.2
GdMg	Cubic ($Pm\bar{3}m$)	90 (120)[g]	15	0.08	4.9

[a][170], [b][171], [c][172], [d][173], [e][174], [f][175], [g][176]

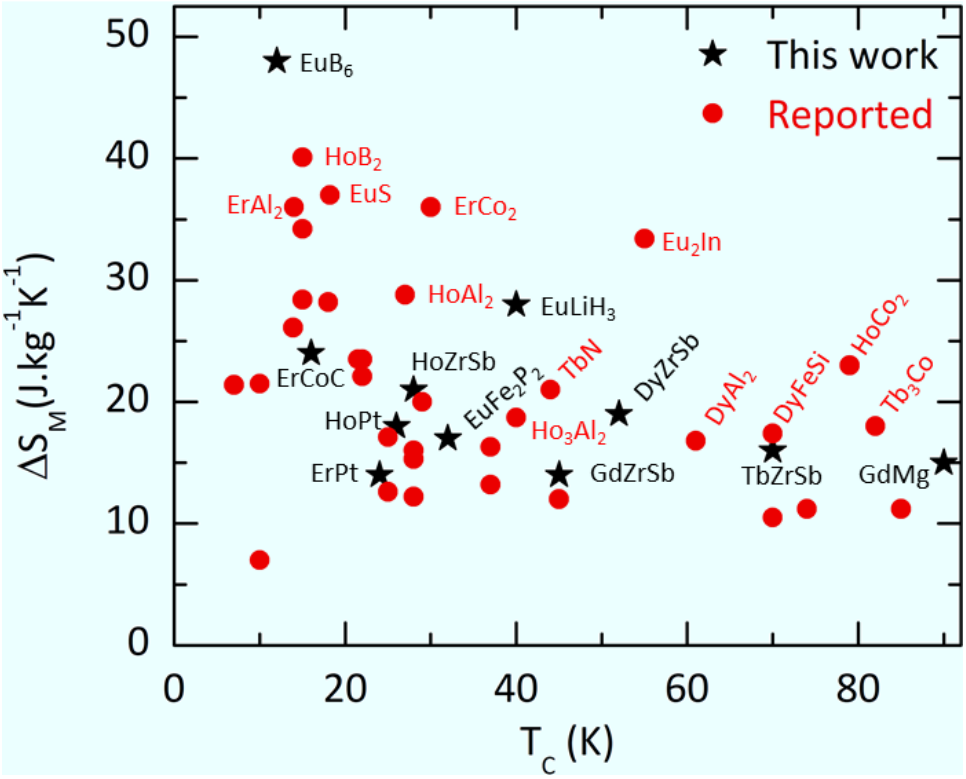

Figure 3.6: **Comparison of newly discovered MCE materials to the known MCE materials.** The magnetocaloric entropy change (ΔS_M) for the predicted compounds in this work (black stars) are plotted alongside some known MCE materials reported in the literature (red circles) in the 10 K - 100 K range for comparison.

ΔS_M, attributed to the relatively larger effective moment associated with Ho compared to the other elements in the series. The predicted compounds are compared with other

materials exhibiting significant magnetocaloric responses in the temperature range of 10 K to 90 K for a field change of 5 T, as illustrated in Fig. 3.6.

3.4 Conclusion

In this work, we exploited LLM's remarkable abilities to summarize texts and create an auto-generated MCE material database from public literature abstracts. Following curation and validation, we trained a neural network with high-quality data to predict good MCE materials. This model efficiently identified unexplored binary and ternary magnetic compounds in the Materials Project database, uncovering promising multiple materials across a wide range of temperatures. We then validated the top predictions using DFT simulations and ASD, discovering new MCE materials for the 10 K to 100 K temperature range. By integrating the results back into the database, we established a feedback loop that enhances the neural network's accuracy. This approach showcases the power of LLMs in mining literature for data-driven material discovery.

3.5 Supplementary Information

3.5.1 Training Data

Figures 3.7(a) and (b) are the frequency distributions of ΔS_M and T_C in the training data. Notably, a large portion of materials in the training dataset have a T_C lower than 100 K. This prevalent characteristic of the dataset is crucial for the model's ability to discover more candidates with a T_C in the desirable range. The frequency histogram of rare-earth elements shown in Fig. 3.7(c) reflects that the majority of reported MCE candidates contain heavy rare-earth metals.

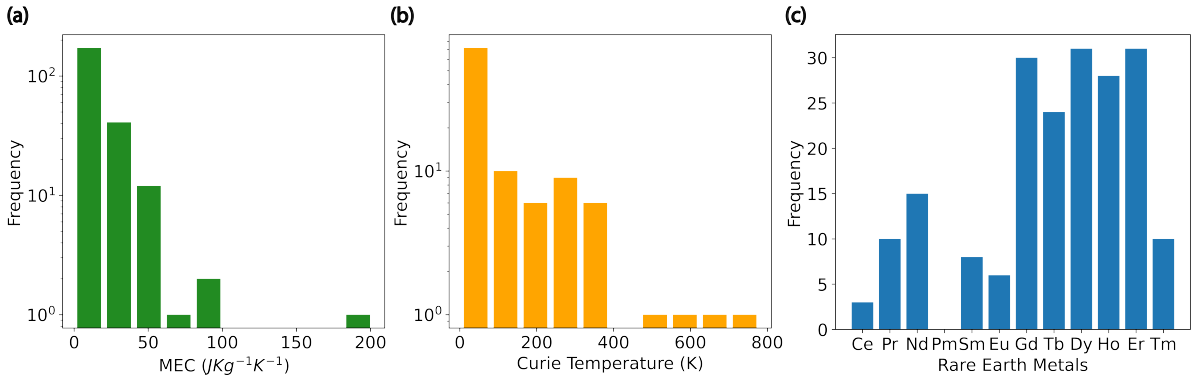

Figure 3.7: Frequency distribution of (a) Magnetic Entropy Change, (b) Curie Temperature, and (c) rare-earth elements in the training data.

3.5.2 Model Training and Evaluation

Figures 3.8 (a) and (b) provide comparisons between the training curves for augmentation factors of 50 and 100, corresponding to datasets with 11,450 and 22,900 training data points, respectively. The model exhibits faster convergence when the training data size is doubled. This compelling evidence underscores the efficacy of the data augmentation method in facilitating the model's learning process. Intriguingly, despite the augmented data not introducing new information, the accelerated convergence suggests that

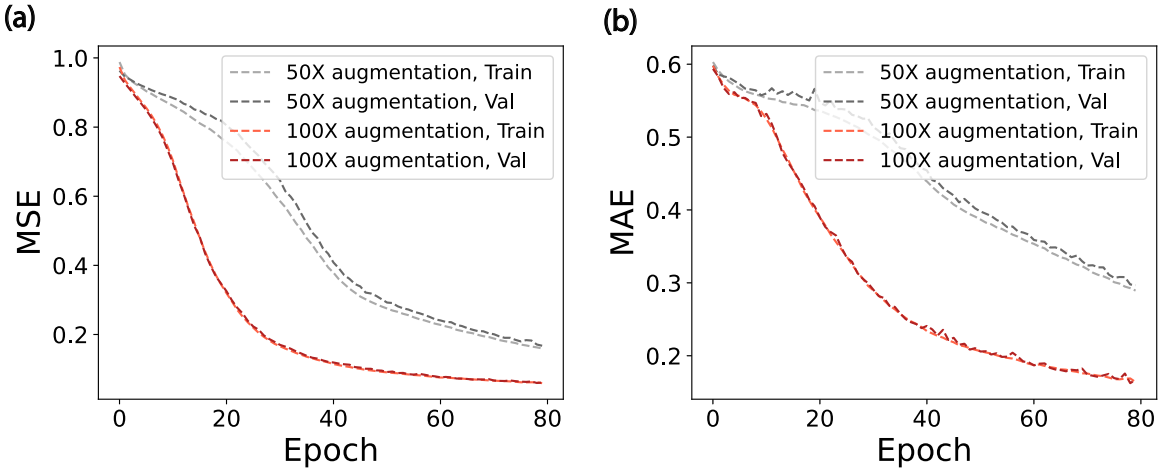

Figure 3.8: Training curves of Mean-Squared Error (MSE) and Mean-Absolute Error (MAE) for augmentation factor of 50 vs. 100.

the increased volume of training instances enhances the model's ability to generalize and capture intricate patterns within the dataset.

Figure 3.9 displays a comprehensive analysis of the grid search results with fixed parameters of 3 layers and 16 neurons, respectively. The training curves offer a visual representation of the model's performance across varying hyperparameter values of the network structure. For each training curve, the x-axis reflects the training epochs and the y-axis illustrates the corresponding values of Mean Squared Error (MSE) or Mean Absolute Error (MAE). It is noteworthy that MSE serves as the employed loss function. Across all configurations, MSE consistently demonstrates smooth convergence without discernible variations, indicating the stability of the model's learning process irrespective of hyperparameter adjustments. The nuanced analysis of MAE, however, provides valuable insights. Among the configurations tested, the networks consist of 3 layers with 16 neurons per layer emerging as the optimal structure.

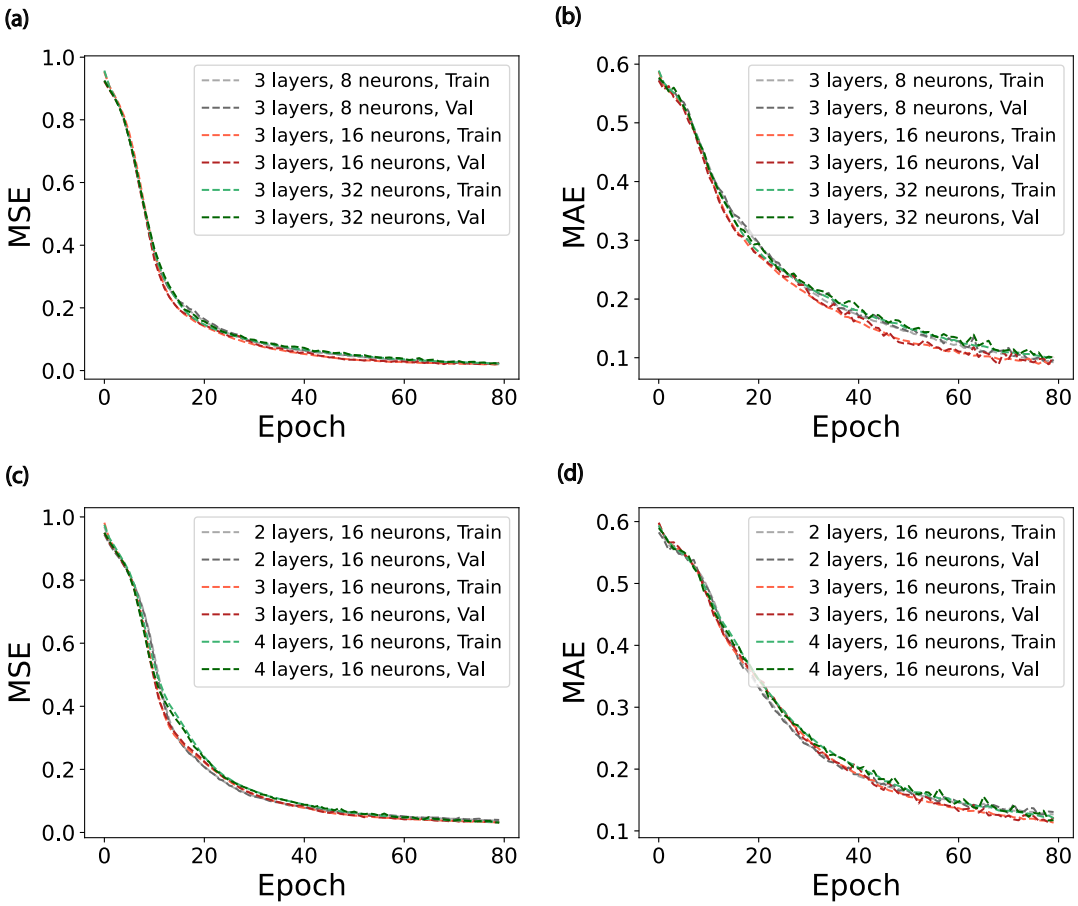

Figure 3.9: Training curves of Mean-Squared Error (MSE) and Mean-Absolute Error (MAE) for grid search.

Data Availability Statement

The data that supports the findings of this study are available within the article and its supplementary information. The auto-generated MCE database is available online at `https://github.com/RunqingYang1996/MCEdatabase`.

Permissions and Attributions

The content of chapter 3 is the result of a collaboration with Runqing Yang, Lokanath Patra, and Bolin Liao, and has previously appeared in the *"Enhancing Magnetocaloric Material Discovery: A Machine Learning Approach Using an Autogenerated Database by Large Language Models"*, (arXiv:2403.02553). It is submitted for publication.

Chapter 4

Summary and Prospects

In this thesis, we combined high-throughput screening with advanced computational methods, state-of-the-art large language models and machine learning techniques to search for promising candidates for future energy applications. Upon comprehensive validation and characterization by first principle calculations, the discovered materials open up possibilities for their potential applications in thermoelectrics devices and magnetic refrigeration. Not only do the results motivate further experimental studies, the framework also allows for efficient screening with more complex materials and offers insights on understanding the performance metric from the descriptors.

The study in Chapter 2 presents a novel approach to identify material candidates with an ultralow intrinsic lattice thermal conductivity using an efficient HTS based on intuitions from the chemical bonding theory. It has been the subject of extensive study to discover crystalline solids with a low thermal conductivity in various applications such as thermoelectric materials, thermal insulation, and thermal barrier coatings. To maximize the thermoelectric figure of merit, for example, it is crucial to minimize the lattice thermal conductivity while maintaining the electrical conductivity, and thus crystalline materials with an ultralow intrinsic thermal conductivity are particularly desirable since

they avoid introducing defects that can interfere electron transport. We reveal a significant correlation between anti-bonding valence bands and low lattice thermal conductivity in crystalline solids using detailed first-principles lattice dynamics analysis. We further utilize the anti-bonding signature of valence bands as an efficient and cost-effective descriptor to conduct high-throughput screening within the Materials Project database to search for stable semiconductors that potentially possess an ultralow intrinsic thermal conductivity. We identify over 600 binary semiconductors with strongly anti-bonding valence bands. From these candidates, we discovered and analyzed a family of semiconductors (XS, where X is an alkali metal) with ultralow thermal conductivity below 1 W/mK at room temperature, which are promising for thermoelectric applications. Our findings provide a deeper understanding of thermal conductivity of crystalline solids at the microscopic chemical bonding level and open a new avenue for discovering materials with intrinsically low thermal conductivity for a broad range of applications.

In Chapter 3, our work demonstrates the potential of combining large language models, machine learning, and *ab initio* simulations to efficiently discover new magnetic refrigerants. Magnetic refrigeration has attracted renewed attention compared to conventional cooling methods due to its solid-state nature, cryogen-free operation, reliability, compact design, and ability to approach Carnot efficiency at cryogenic temperatures. As a result, magnetic refrigeration has found widespread use in space missions, scientific instrumentation, and as an environmentally friendly substitute for vapor compression-based cooling systems. However, most prior research on magnetic refrigeration has concentrated on either extremely low cryogenic temperatures below a few Kelvin or temperatures close to room temperature, whereas many technologically important applications require efficient cooling in the temperature range between 10 K and 100 K, such as hydrogen and methane liquefaction and cryogenic carbon capture. To search for ideal magnetic refrigerants in this temperature range, we develop a computational pipeline to identify

candidates for magnetic refrigeration from a material database generated by extracting more than 6000 entries from open-access literature using a large language model. We then use this database to train a machine learning model that can efficiently predict magnetocaloric properties of materials based on their chemical composition. The magnetocaloric response of the predicted compounds is validated through density functional theory and atomic spin dynamics simulations, allowing for the estimation of magnetic entropy change and adiabatic temperature changes. The pipeline enabled the discovery of 11 new materials exhibiting superior magnetocaloric effect for the target temperature range. Our research showcases the capability of synergizing large language models, machine learning, and *ab initio* simulations for the effective exploration and discovery of novel functional materials.

HTS excels in identifying materials that exhibit exceptional properties within the vast array of structures and compositions, even for challenging task like finding a needle in a haystack. Such discoveries can be surprising, challenging existing knowledge based on smaller datasets. The efficient workflow of descriptors and screening significantly reduces the time and cost devoted for discovering materials with desired properties. This accelerates the process to attain more profound insights into material properties across a range of applications. With the increasing accessibility and growth of datasets, the exploration of data using modern AI techniques is becoming increasingly inevitable. This trend is expected to yield new insights in the relationships between structure, chemistry, and properties. Undoubtedly, these advancements will propel materials science towards a more data-driven paradigm, fundamentally transforming the traditional methods of discovering and developing new materials.

Bibliography

[1] S. Chu, Y. Cui, and N. Liu, *The path towards sustainable energy*, *Nature materials* **16** (2017), no. 1 16–22.

[2] M. A. Green, *The path to 25% silicon solar cell efficiency: History of silicon cell evolution*, *Progress in photovoltaics: research and applications* **17** (2009), no. 3 183–189.

[3] L. C. Andreani, A. Bozzola, P. Kowalczewski, M. Liscidini, and L. Redorici, *Silicon solar cells: toward the efficiency limits*, *Advances in Physics: X* **4** (2019), no. 1 1548305.

[4] Y. Rong, Y. Hu, A. Mei, H. Tan, M. I. Saidaminov, S. I. Seok, M. D. McGehee, E. H. Sargent, and H. Han, *Challenges for commercializing perovskite solar cells*, *Science* **361** (2018), no. 6408 eaat8235.

[5] Y.-W. Su, S.-C. Lan, and K.-H. Wei, *Organic photovoltaics*, *Materials Today* **15** (2012), no. 12 554–562.

[6] P. K. Nayak, S. Mahesh, H. J. Snaith, and D. Cahen, *Photovoltaic solar cell technologies: analysing the state of the art*, *Nature Reviews Materials* **4** (2019), no. 4 269–285.

[7] R. Marom, S. F. Amalraj, N. Leifer, D. Jacob, and D. Aurbach, *A review of advanced and practical lithium battery materials*, *Journal of Materials Chemistry* **21** (2011), no. 27 9938–9954.

[8] L. Sun, Y. Liu, R. Shao, J. Wu, R. Jiang, and Z. Jin, *Recent progress and future perspective on practical silicon anode-based lithium ion batteries*, *Energy Storage Materials* **46** (2022) 482–502.

[9] Y. You, H. Celio, J. Li, A. Dolocan, and A. Manthiram, *Modified high-nickel cathodes with stable surface chemistry against ambient air for lithium-ion batteries*, *Angewandte Chemie International Edition* **57** (2018), no. 22 6480–6485.

[10] J.-Y. Hwang, S.-T. Myung, and Y.-K. Sun, *Sodium-ion batteries: present and future*, *Chemical Society Reviews* **46** (2017), no. 12 3529–3614.

[11] V. Keppens, B. Sales, D. Mandrus, B. Chakoumakos, and C. Laermans, *When does a crystal conduct heat like a glass?*, *Phil. Mag. Lett.* **80** (Dec., 2000) 807–812.

[12] G. S. Nolas, G. A. Slack, and S. B. Schujman, *Chapter 6 semiconductor clathrates: A phonon glass electron crystal material with potential for thermoelectric applications*, in *Recent Trends in Thermoelectric Materials Research I*, vol. 69, pp. 255–300. Elsevier, 2001.

[13] M. Beekman, D. T. Morelli, and G. S. Nolas, *Better thermoelectrics through glass-like crystals*, *Nat. Mater.* **14** (Nov., 2015) 1182–1185.

[14] B. Liao and G. Chen, *Nanocomposites for thermoelectrics and thermal engineering*, *MRS Bull.* **40** (Sept., 2015) 746–752.

[15] A. J. Minnich, M. S. Dresselhaus, Z. F. Ren, and G. Chen, *Bulk nanostructured thermoelectric materials: Current research and future prospects*, *Energy Environ. Sci.* **2** (2009), no. 5 466.

[16] K. Biswas, J. He, I. D. Blum, C.-I. Wu, T. P. Hogan, D. N. Seidman, V. P. Dravid, and M. G. Kanatzidis, *High-performance bulk thermoelectrics with all-scale hierarchical architectures*, *Nature* **489** (Sept., 2012) 414–418.

[17] B. Poudel, Q. Hao, Y. Ma, Y. Lan, A. Minnich, B. Yu, X. Yan, D. Wang, A. Muto, D. Vashaee, X. Chen, J. Liu, M. S. Dresselhaus, G. Chen, and Z. Ren, *High-thermoelectric performance of nanostructured bismuth antimony Telluride bulk alloys*, *Science* **320** (May, 2008) 634–638.

[18] J. Ravichandran, A. K. Yadav, R. Cheaito, P. B. Rossen, A. Soukiassian, S. J. Suresha, J. C. Duda, B. M. Foley, C.-H. Lee, Y. Zhu, A. W. Lichtenberger, J. E. Moore, D. A. Muller, D. G. Schlom, P. E. Hopkins, A. Majumdar, R. Ramesh, and M. A. Zurbuchen, *Crossover from incoherent to coherent phonon scattering in epitaxial oxide superlattices*, *Nat. Mater.* **13** (Dec., 2013) 168–172.

[19] Z. Chen, X. Zhang, and Y. Pei, *Manipulation of phonon transport in thermoelectrics*, *Adv. Mater.* **30** (Feb., 2018) 1705617.

[20] X. Qian, J. Zhou, and G. Chen, *Phonon-engineered extreme thermal conductivity materials*, *Nat. Mater.* **20** (Mar., 2021) 1188–1202.

[21] J. Mao, Z. Liu, J. Zhou, H. Zhu, Q. Zhang, G. Chen, and Z. Ren, *Advances in thermoelectrics*, *Adv. Phys.* **67** (Apr., 2018) 69–147.

[22] G. K. Madsen, *Automated search for new thermoelectric materials: the case of liznsb*, *Journal of the American Chemical Society* **128** (2006), no. 37 12140–12146.

[23] T. Wang, Y. Xiong, Y. Wang, P. Qiu, Q. Song, K. Zhao, J. Yang, J. Xiao, X. Shi, and L. Chen, *Cu_3ErTe_3: A new promising thermoelectric material predicated by high-throughput screening*, *Mater. Today Phys.* **12** (Mar., 2020) 100180.

[24] J. Carrete, W. Li, N. Mingo, S. Wang, and S. Curtarolo, *Finding unprecedentedly low-thermal-conductivity half-heusler semiconductors via high-throughput materials modeling*, *Phys. Rev. X* **4** (Feb., 2014) 011019.

[25] V. K. Pecharsky and K. A. Gschneidner Jr, *Magnetocaloric effect and magnetic refrigeration*, *Journal of Magnetism and Magnetic Materials* **200** (1999), no. 1-3 44–56.

[26] C. Liu, X. Wang, W. Cai, and H. Su, *Prediction of magnetocaloric properties of fe-based amorphous alloys based on interpretable machine learning*, *Journal of Non-Crystalline Solids* **625** (2024) 122749.

[27] A. Alahmer, M. Al-Amayreh, A. O. Mostafa, M. Al-Dabbas, and H. Rezk, *Magnetic refrigeration design technologies: State of the art and general perspectives*, *Energies* **14** (2021), no. 15 4662.

[28] G. Hautier, A. Jain, H. Chen, C. Moore, S. P. Ong, and G. Ceder, *Novel mixed polyanions lithium-ion battery cathode materials predicted by high-throughput ab initio computations*, *Journal of Materials Chemistry* **21** (2011), no. 43 17147–17153.

[29] P. B. d. Castro, K. Terashima, T. D. Yamamoto, Z. Hou, S. Iwasaki, R. Matsumoto, S. Adachi, Y. Saito, P. Song, H. Takeya, *et. al.*, *Machine-learning-guided discovery of the gigantic magnetocaloric effect in HoB_2 near the hydrogen liquefaction temperature*, *NPG Asia Materials* **12** (2020), no. 1 35.

[30] N. J. Szymanski, B. Rendy, Y. Fei, R. E. Kumar, T. He, D. Milsted, M. J. McDermott, M. Gallant, E. D. Cubuk, A. Merchant, *et. al.*, *An autonomous laboratory for the accelerated synthesis of novel materials*, *Nature* **624** (2023), no. 7990 86–91.

[31] A. Merchant, S. Batzner, S. S. Schoenholz, M. Aykol, G. Cheon, and E. D. Cubuk, *Scaling deep learning for materials discovery*, *Nature* **624** (2023), no. 7990 80–85.

[32] K. M. Jablonka, Q. Ai, A. Al-Feghali, S. Badhwar, J. D. Bocarsly, A. M. Bran, S. Bringuier, L. C. Brinson, K. Choudhary, D. Circi, *et. al.*, *14 examples of how LLMs can transform materials science and chemistry: a reflection on a large language model hackathon*, *Digital Discovery* **2** (2023), no. 5 1233–1250.

[33] G. J. Snyder and E. S. Toberer, *Complex thermoelectric materials*, *Nat. Mater.* **7** (Feb., 2008) 105–114.

[34] D. Zhao and G. Tan, *A review of thermoelectric cooling: Materials, modeling and applications*, *Applied thermal engineering* **66** (2014), no. 1-2 15–24.

[35] B.-T. Wang and P. Zhang, *Phonon spectrum and bonding properties of bi2se3: Role of strong spin-orbit interaction*, *Applied Physics Letters* **100** (2012), no. 8.

[36] H. Mamur, M. Bhuiyan, F. Korkmaz, and M. Nil, *A review on bismuth telluride (bi2te3) nanostructure for thermoelectric applications*, *Renewable and Sustainable Energy Reviews* **82** (2018) 4159–4169.

[37] Y. Yu, M. Cagnoni, O. Cojocaru-Mirédin, and M. Wuttig, *Chalcogenide thermoelectrics empowered by an unconventional bonding mechanism*, *Advanced Functional Materials* **30** (2020), no. 8 1904862.

[38] Z.-G. Chen, X. Shi, L.-D. Zhao, and J. Zou, *High-performance snse thermoelectric materials: Progress and future challenge*, *Progress in Materials Science* **97** (2018) 283–346.

[39] J. P. Heremans, V. Jovovic, E. S. Toberer, A. Saramat, K. Kurosaki, A. Charoenphakdee, S. Yamanaka, and G. J. Snyder, *Enhancement of thermoelectric efficiency in PbTe by distortion of the electronic density of states*, *Science* **321** (July, 2008) 554–557.

[40] O. Delaire, J. Ma, K. Marty, A. F. May, M. A. McGuire, M.-H. Du, D. J. Singh, A. Podlesnyak, G. Ehlers, M. D. Lumsden, and B. C. Sales, *Giant anharmonic phonon scattering in PbTe*, *Nat. Mater.* **10** (June, 2011) 614–619.

[41] Y. Pei, J. Lensch-Falk, E. S. Toberer, D. L. Medlin, and G. J. Snyder, *High thermoelectric performance in PbTe due to large nanoscale Ag2Te precipitates and la doping*, *Adv. Funct. Mater.* **21** (Nov., 2010) 241–249.

[42] G. A. Akhmedova and D. S. Abdinov, *Effect of thallium doping on the thermal conductivity of PbTe single crystals*, *Inorg. Mater.* **45** (July, 2009) 854–858.

[43] G. Nolas, D. Morelli, and T. M. Tritt, *Skutterudites: A phonon-glass-electron crystal approach to advanced thermoelectric energy conversion applications*, *Annual Review of Materials Science* **29** (1999), no. 1 89–116.

[44] T. Zhu, C. Fu, H. Xie, Y. Liu, and X. Zhao, *High efficiency half-heusler thermoelectric materials for energy harvesting*, *Advanced Energy Materials* **5** (2015), no. 19 1500588.

[45] E. K. Lee, L. Yin, Y. Lee, J. W. Lee, S. J. Lee, J. Lee, S. N. Cha, D. Whang, G. S. Hwang, K. Hippalgaonkar, *et. al.*, *Large thermoelectric figure-of-merits from sige nanowires by simultaneously measuring electrical and thermal transport properties*, *Nano letters* **12** (2012), no. 6 2918–2923.

[46] C. Fu, Y. Sun, and C. Felser, *Topological thermoelectrics*, *APL Materials* **8** (2020), no. 4.

[47] T. Schumann, M. Goyal, H. Kim, and S. Stemmer, *Molecular beam epitaxy of cd3as2 on a iii-v substrate*, *APL Materials* **4** (2016), no. 12.

[48] D. Vu, W. Zhang, C. Şahin, M. E. Flatté, N. Trivedi, and J. P. Heremans, *Thermal chiral anomaly in the magnetic-field-induced ideal weyl phase of bi1- x sb x*, *Nature materials* **20** (2021), no. 11 1525–1531.

[49] N. Toshima, *Recent progress of organic and hybrid thermoelectric materials*, *Synthetic Metals* **225** (2017) 3–21.

[50] B. Yu, M. Liu, P. W. Egolf, and A. Kitanovski, *A review of magnetic refrigerator and heat pump prototypes built before the year 2010*, *International Journal of refrigeration* **33** (2010), no. 6 1029–1060.

[51] N. A. Zarkevich and V. I. Zverev, *Viable materials with a giant magnetocaloric effect*, *Crystals* **10** (2020), no. 9 815.

[52] N. Terada, K. Terashima, P. B. De Castro, C. V. Colin, H. Mamiya, T. D. Yamamoto, H. Takeya, O. Sakai, Y. Takano, and H. Kitazawa, *Relationship between magnetic ordering and gigantic magnetocaloric effect in HoB_2 studied by neutron diffraction experiment*, *Physical Review B* **102** (2020), no. 9 094435.

[53] H. Wada, Y. Tanabe, M. Shiga, H. Sugawara, and H. Sato, *Magnetocaloric effects of Laves phase $Er(Co_{1-x}Ni_x)_2$ compounds*, *Journal of alloys and compounds* **316** (2001), no. 1-2 245–249.

[54] P. Von Ranke, V. Pecharsky, K. Gschneidner, *et. al.*, *Influence of the crystalline electrical field on the magnetocaloric effect of $DyAl_2$, $ErAl_2$, and $DyNi_2$*, *Physical review B* **58** (1998), no. 18 12110.

[55] P. Von Ranke, N. De Oliveira, E. Plaza, V. De Sousa, B. Alho, A. M. G. Carvalho, S. Gama, and M. Reis, *The giant anisotropic magnetocaloric effect in $DyAl_2$*, *Journal of Applied Physics* **104** (2008), no. 9.

[56] C. M. Acosta, E. Ogoshi, J. A. Souza, and G. M. Dalpian, *Machine learning study of the magnetic ordering in 2d materials*, *ACS Applied Materials & Interfaces* **14** (2022), no. 7 9418–9432.

[57] T. Numazawa, K. Kamiya, T. Utaki, and K. Matsumoto, *Magnetic refrigerator for hydrogen liquefaction*, *Cryogenics* **62** (2014) 185–192.

[58] Y. Kim, I. Park, and S. Jeong, *Experimental investigation of two-stage active magnetic regenerative refrigerator operating between 77K and 20K*, *Cryogenics* **57** (2013) 113–121.

[59] M.-H. Phan and S.-C. Yu, *Review of the magnetocaloric effect in manganite materials*, *Journal of Magnetism and Magnetic Materials* **308** (2007), no. 2 325–340.

[60] S. Arapan, P. Nieves, and S. Cuesta-López, *A high-throughput exploration of magnetic materials by using structure predicting methods*, *Journal of Applied Physics* **123** (2018), no. 8.

[61] Y. Zhang and X. Xu, *Machine learning the magnetocaloric effect in manganites from compositions and structural parameters*, *AIP Advances* **10** (2020), no. 3.

[62] A. G. Kusne, T. Gao, A. Mehta, L. Ke, M. C. Nguyen, K.-M. Ho, V. Antropov, C.-Z. Wang, M. J. Kramer, C. Long, *et. al.*, *On-the-fly machine-learning for high-throughput experiments: search for rare-earth-free permanent magnets*, *Scientific Reports* **4** (2014), no. 1 6367.

[63] D. Silva, J. Ventura, and J. Araújo, *Predicting the performance of magnetocaloric systems using machine learning regressors*, *Energy and AI* **2** (2020) 100030.

[64] J. D. Bocarsly, E. E. Levin, C. A. Garcia, K. Schwennicke, S. D. Wilson, and R. Seshadri, *A simple computational proxy for screening magnetocaloric compounds*, *Chem. Mater.* **29** (2017), no. 4 1613–1622.

[65] C. J. Court, A. Jain, and J. M. Cole, *Inverse design of materials that exhibit the magnetocaloric effect by text-mining of the scientific literature and generative deep learning*, *Chemistry of Materials* **33** (2021), no. 18 7217–7231.

[66] J. Lai, A. Bolyachkin, N. Terada, S. Dieb, X. Tang, T. Ohkubo, H. Sepehri-Amin, and K. Hono, *Machine learning assisted development of Fe_2P-type magnetocaloric compounds for cryogenic applications*, *Acta Materialia* **232** (2022) 117942.

[67] D. R. Clarke and S. R. Phillpot, *Thermal barrier coating materials*, *Mater. Today* **8** (June, 2005) 22–29.

[68] T. Takabatake, K. Suekuni, T. Nakayama, and E. Kaneshita, *Phonon-glass electron-crystal thermoelectric clathrates: Experiments and theory*, *Rev. Mod. Phys.* **86** (June, 2014) 669–716.

[69] B. Sun, S. Niu, R. P. Hermann, J. Moon, N. Shulumba, K. Page, B. Zhao, A. S. Thind, K. Mahalingam, J. Milam-Guerrero, R. Haiges, M. Mecklenburg, B. C. Melot, Y.-D. Jho, B. M. Howe, R. Mishra, A. Alatas, B. Winn, M. E. Manley, J. Ravichandran, and A. J. Minnich, *High frequency atomic tunneling yields ultralow and glass-like thermal conductivity in chalcogenide single crystals*, *Nat. Commun.* **11** (Nov., 2020) 6039.

[70] J. Ma, O. Delaire, A. F. May, C. E. Carlton, M. A. McGuire, L. H. VanBebber, D. L. Abernathy, G. Ehlers, T. Hong, A. Huq, W. Tian, V. M. Keppens, Y. Shao-Horn, and B. C. Sales, *Glass-like phonon scattering from a spontaneous nanostructure in $AgSbTe_2$*, *Nat. Nanotechnol.* **8** (June, 2013) 445–451.

[71] L.-D. Zhao, S.-H. Lo, Y. Zhang, H. Sun, G. Tan, C. Uher, C. Wolverton, V. P. Dravid, and M. G. Kanatzidis, *Ultralow thermal conductivity and high thermoelectric figure of merit in SnSe crystals*, *Nature* **508** (Apr., 2014) 373–377.

[72] X. Lu, D. T. Morelli, Y. Xia, F. Zhou, V. Ozolins, H. Chi, X. Zhou, and C. Uher, *High performance thermoelectricity in earth-abundant compounds based on natural mineral tetrahedrites*, *Adv. Energy Mater.* **3** (Oct., 2012) 342–348.

[73] J. He, Y. Xia, W. Lin, K. Pal, Y. Zhu, M. G. Kanatzidis, and C. Wolverton, *Accelerated discovery and design of ultralow lattice thermal conductivity materials using chemical bonding principles*, *Adv. Funct. Mater.* **32** (Dec., 2021) 2108532.

[74] C. Chiritescu, D. G. Cahill, N. Nguyen, D. Johnson, A. Bodapati, P. Keblinski, and P. Zschack, *Ultralow thermal conductivity in disordered, layered WSe_2 crystals*, *Science* **315** (Jan., 2007) 351–353.

[75] R. M. Costescu, D. G. Cahill, F. H. Fabreguette, Z. A. Sechrist, and S. M. George, *Ultra-low thermal conductivity in W/Al_2O_3 nanolaminates*, *Science* **303** (Feb., 2004) 989–990.

[76] H. Liu, X. Shi, F. Xu, L. Zhang, W. Zhang, L. Chen, Q. Li, C. Uher, T. Day, and G. J. Snyder, *Copper ion liquid-like thermoelectrics*, *Nat. Mater.* **11** (Mar., 2012) 422–425.

[77] S. Roychowdhury, M. K. Jana, J. Pan, S. N. Guin, D. Sanyal, U. V. Waghmare, and K. Biswas, *Soft phonon modes leading to ultralow thermal conductivity and high thermoelectric performance in AgCuTe*, *Angew. Chem.* **130** (Mar., 2018) 4107–4111.

[78] B. Li, H. Wang, Y. Kawakita, Q. Zhang, M. Feygenson, H. L. Yu, D. Wu, K. Ohara, T. Kikuchi, K. Shibata, T. Yamada, X. K. Ning, Y. Chen, J. Q. He, D. Vaknin, R. Q. Wu, K. Nakajima, and M. G. Kanatzidis, *Liquid-like thermal conduction in intercalated layered crystalline solids*, *Nat. Mater.* **17** (Jan., 2018) 226–230.

[79] H. Xie, S. Hao, J. Bao, T. J. Slade, G. J. Snyder, C. Wolverton, and M. G. Kanatzidis, *All-inorganic halide perovskites as potential thermoelectric materials: Dynamic cation off-centering induces ultralow thermal conductivity*, *JACS* **142** (Apr., 2020) 9553–9563.

[80] D. Sarkar, T. Ghosh, S. Roychowdhury, R. Arora, S. Sajan, G. Sheet, U. V. Waghmare, and K. Biswas, *Ferroelectric instability induced ultralow thermal conductivity and high thermoelectric performance in rhombohedral p-type GeSe crystal*, *JACS* **142** (June, 2020) 12237–12244.

[81] T. Tritt, *Recent trends in thermoelectric materials research: Part three*. Elsevier, 2001.

[82] T. Tadano, Y. Gohda, and S. Tsuneyuki, *Impact of rattlers on thermal conductivity of a thermoelectric clathrate: A first-principles study*, *Phys. Rev. Lett.* **114** (Mar., 2015) 095501.

[83] H. Lin, G. Tan, J.-N. Shen, S. Hao, L.-M. Wu, N. Calta, C. Malliakas, S. Wang, C. Uher, C. Wolverton, and M. G. Kanatzidis, *Concerted rattling in $CsAg_5Te_3$ leading to ultralow thermal conductivity and high thermoelectric performance*, *Angew. Chem. Int. Ed.* **55** (Aug., 2016) 11431–11436.

[84] J. He, M. Amsler, Y. Xia, S. S. Naghavi, V. I. Hegde, S. Hao, S. Goedecker, V. Ozoliņš, and C. Wolverton, *Ultralow thermal conductivity in full heusler semiconductors*, *Phys. Rev. Lett.* **117** (July, 2016) 046602.

[85] M. D. Nielsen, V. Ozolins, and J. P. Heremans, *Lone pair electrons minimize lattice thermal conductivity*, *Energy Environ. Sci.* **6** (2013), no. 2 570–578.

[86] E. J. Skoug and D. T. Morelli, *Role of lone-pair electrons in producing minimum thermal conductivity in nitrogen-group chalcogenide compounds*, *Phys. Rev. Lett.* **107** (Nov., 2011) 235901.

[87] M. Dutta, K. Pal, U. V. Waghmare, and K. Biswas, *Bonding heterogeneity and lone pair induced anharmonicity resulted in ultralow thermal conductivity and promising thermoelectric properties in n-type $AgPbBiSe_3$*, *Chem. Sci.* **10** (2019), no. 18 4905–4913.

[88] G. S. Nolas, J. Sharp, and H. J. Goldsmid, *Thermoelectrics: Basic Principles and New Materials Developments*, vol. 45. Springer Berlin Heidelberg, 2001.

[89] Z. Chen, Z. Jian, W. Li, Y. Chang, B. Ge, R. Hanus, J. Yang, Y. Chen, M. Huang, G. J. Snyder, and Y. Pei, *Lattice dislocations enhancing thermoelectric PbTe in addition to band convergence*, *Adv. Mater.* **29** (Apr., 2017) 1606768.

[90] L.-D. Zhao, G. Tan, S. Hao, J. He, Y. Pei, H. Chi, H. Wang, S. Gong, H. Xu, V. P. Dravid, C. Uher, G. J. Snyder, C. Wolverton, and M. G. Kanatzidis, *Ultrahigh power factor and thermoelectric performance in hole-doped single-crystal SnSe*, *Science* **351** (Jan., 2016) 141–144.

[91] J. Li, X. Zhang, Z. Chen, S. Lin, W. Li, J. Shen, I. T. Witting, A. Faghaninia, Y. Chen, A. Jain, L. Chen, G. J. Snyder, and Y. Pei, *Low-symmetry rhombohedral GeTe thermoelectrics*, *Joule* **2** (May, 2018) 976–987.

[92] W.-D. Liu, D.-Z. Wang, Q. Liu, W. Zhou, Z. Shao, and Z.-G. Chen, *High-Performance GeTe-Based thermoelectrics: From materials to devices*, *Adv. Energy Mater.* **10** (Apr., 2020) 2000367.

[93] S. Lee, K. Esfarjani, T. Luo, J. Zhou, Z. Tian, and G. Chen, *Resonant bonding leads to low lattice thermal conductivity*, *Nat. Commun.* **5** (Apr., 2014) 3525.

[94] T. Yue, Y. Zhao, J. Ni, S. Meng, and Z. Dai, *Strong quartic anharmonicity, ultralow thermal conductivity, high band degeneracy and good thermoelectric performance in Na_2TlSb*, *npj Comput. Mater.* **9** (Feb., 2023) 17.

[95] J. Zhang, D. Ishikawa, M. M. Koza, E. Nishibori, L. Song, A. Q. R. Baron, and B. B. Iversen, *Dynamic lone pair expression as chemical bonding origin of giant phonon anharmonicity in thermoelectric InTe*, *Angew. Chem.* **135** (Feb., 2023) e202218458.

[96] W. Wan, Y. Ge, and Y. Liu, *Strong phonon anharmonicity and low thermal conductivity of monolayer tin oxides driven by lone-pair electrons*, *Appl. Phys. Lett.* **114** (Jan., 2019) 031901.

[97] J. Y. Kim, J.-W. Lee, H. S. Jung, H. Shin, and N.-G. Park, *High-efficiency perovskite solar cells*, *Chem. Rev.* **120** (July, 2020) 7867–7918.

[98] H. Ma, C. Li, Y. Ma, H. Wang, Z. W. Rouse, Z. Zhang, C. Slebodnick, A. Alatas, S. P. Baker, J. J. Urban, and Z. Tian, *Supercompliant and soft $(CH_3NH_3)_3Bi_2I_9$ crystal with ultralow thermal conductivity*, *Phys. Rev. Lett.* **123** (Oct., 2019) 155901.

[99] D. H. Fabini, R. Seshadri, and M. G. Kanatzidis, *The underappreciated lone pair in halide perovskites underpins their unusual properties*, *MRS Bull.* **45** (June, 2020) 467–477.

[100] T. Kim, S. Park, V. Iyer, B. Shaheen, U. Choudhry, Q. Jiang, G. Eichman, R. Gnabasik, K. Kelley, B. Lawrie, K. Zhu, and B. Liao, *Mapping the pathways of photo-induced ion migration in organic-inorganic hybrid halide perovskites*, *Nat. Commun.* **14** (Apr., 2023) 1846.

[101] J. Li, W. Hu, and J. Yang, *High-throughput screening of rattling-induced ultralow lattice thermal conductivity in semiconductors*, *JACS* **144** (Mar., 2022) 4448–4456.

[102] A. Rodriguez, C. Lin, H. Yang, M. Al-Fahdi, C. Shen, K. Choudhary, Y. Zhao, J. Hu, B. Cao, H. Zhang, and M. Hu, *Million-scale data integrated deep neural network for phonon properties of heuslers spanning the periodic table*, *npj Comput. Mater.* **9** (Feb., 2023) 20.

[103] K. Yuan, X. Zhang, Z. Chang, D. Tang, and M. Hu, *Antibonding induced anharmonicity leading to ultralow lattice thermal conductivity and extraordinary thermoelectric performance in CsK_2X (X = Sb, Bi)*, *J. Mater. Chem. C* **10** (2022), no. 42 15822–15832.

[104] A. Das, K. Pal, P. Acharyya, S. Das, K. Maji, and K. Biswas, *Strong antibonding I (p)–Cu (d) states lead to intrinsically low thermal conductivity in $CuBiI_4$*, *JACS* **145** (Jan., 2023) 1349–1358.

[105] J. Zhang, H. Jiang, X. Xia, Y. Gao, and Z. Huang, *Antibonding p-d and s-p hybridization induce the optimization of thermal and thermoelectric performance of $MGeTe_3$ (M = In and Sb)*, *ACS Appl. Energy Mater.* **5** (Dec., 2022) 15566–15577.

[106] R. Dronskowski and P. E. Bloechl, *Crystal orbital Hamilton populations (COHP): Energy-resolved visualization of chemical bonding in solids based on density-functional calculations*, *The Journal of Physical Chemistry* **97** (Aug., 1993) 8617–8624.

[107] V. L. Deringer, A. L. Tchougréeff, and R. Dronskowski, *Crystal orbital Hamilton population (COHP) analysis as projected from plane-wave basis sets*, *J. Phys. Chem. A* **115** (May, 2011) 5461–5466.

[108] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. A. Persson, *Commentary: The materials project: A materials genome approach to accelerating materials innovation*, *APL Mater.* **1** (July, 2013) 011002.

[109] G. Kresse and J. Furthmüller, *Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set*, *Phys. Rev. B* **54** (Oct., 1996) 11169–11186.

[110] G. Kresse and J. Furthmüller, *Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set*, *Comput. Mater. Sci.* **6** (July, 1996) 15–50.

[111] P. E. Blöchl, *Projector augmented-wave method*, *Phys. Rev. B* **50** (Dec., 1994) 17953–17979.

[112] J. P. Perdew, K. Burke, and M. Ernzerhof, *Generalized gradient approximation made simple*, *Phys. Rev. Lett.* **77** (Oct., 1996) 3865–3868.

[113] D. A. Broido, M. Malorny, G. Birner, N. Mingo, and D. A. Stewart, *Intrinsic lattice thermal conductivity of semiconductors from first principles*, *Appl. Phys. Lett.* **91** (Dec., 2007) 231922.

[114] A. Togo and I. Tanaka, *First principles phonon calculations in materials science*, *Scripta Mater.* **108** (Nov., 2015) 1–5.

[115] W. Li, J. Carrete, N. A. Katcho, and N. Mingo, *ShengBTE: A solver of the Boltzmann transport equation for phonons*, *Comput. Phys. Commun.* **185** (June, 2014) 1747–1758.

[116] J. A. Pople and D. L. Beveridge, *Molecular orbital theory*, *C0., NY* (1970).

[117] W. J. Hehre, *Ab initio molecular orbital theory*, *Accounts Chem. Res.* **9** (Nov., 1976) 399–406.

[118] U. V. Waghmare, N. A. Spaldin, H. C. Kandpal, and R. Seshadri, *First-principles indicators of metallicity and cation off-centricity in the IV-VI rocksalt chalcogenides of divalent Ge, Sn, and Pb*, *Phys. Rev. B* **67** (Mar., 2003) 125111.

[119] Z. Tian, J. Garg, K. Esfarjani, T. Shiga, J. Shiomi, and G. Chen, *Phonon conduction in PbSe, PbTe, and $PbTe_{1-x}Se_x$ from first-principles calculations*, *Physical Review B* **85** (2012), no. 18 184303.

[120] A. Jain and A. J. McGaughey, *Effect of exchange–correlation on first-principles-driven lattice thermal conductivity predictions of crystalline silicon*, *Comput. Mater. Sci.* **110** (Dec., 2015) 115–120.

[121] E. H. Jung, N. J. Jeon, E. Y. Park, C. S. Moon, T. J. Shin, T.-Y. Yang, J. H. Noh, and J. Seo, *Efficient, stable and scalable perovskite solar cells using poly(3-hexylthiophene)*, *Nature* **567** (Mar., 2019) 511–515.

[122] Z. Li, B. Li, X. Wu, S. A. Sheppard, S. Zhang, D. Gao, N. J. Long, and Z. Zhu, *Organometallic-functionalized interfaces for highly efficient inverted perovskite solar cells*, *Science* **376** (Apr., 2022) 416–420.

[123] M. A. Haque, S. Kee, D. R. Villalva, W.-L. Ong, and D. Baran, *Halide perovskites: Thermal transport and prospects for thermoelectricity*, *Adv. Sci.* **7** (Apr., 2020) 1903389.

[124] Y. Wang, R. Lin, P. Zhu, Q. Zheng, Q. Wang, D. Li, and J. Zhu, *Cation dynamics governed thermal properties of lead halide perovskite nanowires*, *Nano Lett.* **18** (Apr., 2018) 2772–2779.

[125] R. Seshadri, G. Baldinozzi, C. Felser, and W. Tremel, *Visualizing electronic structure changes across an antiferroelectric phase transition: Pb_2MgWO_6*, *J. Mater. Chem.* **9** (1999), no. 10 2463–2466.

[126] J.-M. Raulot, G. Baldinozzi, R. Seshadri, and P. Cortona, *An ab-initio study of the rôle of lone pairs in the structure and insulator–metal transition in SnO and PbO*, *Solid State Sci.* **4** (Apr., 2002) 467–474.

[127] M. W. Stoltzfus, P. M. Woodward, R. Seshadri, J.-H. Klepeis, and B. Bursten, *Structure and bonding in $SnWO_4$, $PbWO_4$, and $BiVO_4$: lone pairs vs inert pairs*, *Inorg. Chem.* **46** (Apr., 2007) 3839–3850.

[128] O. Hellman, P. Steneteg, I. A. Abrikosov, and S. I. Simak, *Temperature dependent effective potential method for accurate free energy calculations of solids*, *Phys. Rev. B* **87** (Mar., 2013) 104111.

[129] T. Lanigan-Atkins, X. He, M. J. Krogstad, D. M. Pajerowski, D. L. Abernathy, G. N. M. N. Xu, Z. Xu, D.-Y. Chung, M. G. Kanatzidis, S. Rosenkranz, R. Osborn, and O. Delaire, *Two-dimensional overdamped fluctuations of the soft perovskite lattice in $CsPbBr_3$*, *Nat. Mater.* **20** (Mar., 2021) 977–983.

[130] C. Gehrmann and D. A. Egger, *Dynamic shortening of disorder potentials in anharmonic halide perovskites*, *Nat. Commun.* **10** (July, 2019) 3141.

[131] Y. Li, X. Jin, T. Cui, Q. Zhuang, D. Zhang, X. Meng, K. Bao, B. Liu, and Q. Zhou, *Unexpected stable stoichiometries and superconductivity of potassium-rich sulfides*, *RSC Adv.* **7** (2017), no. 71 44884–44889.

[132] J. Sangster and A. Pelton, *The KS (potassium-sulfur) system*, *J. Phase Equilib.* **18** (1997), no. 1 82.

[133] P. Böttcher, J. Getzschmann, and R. Keller, *Zur kenntnis der dialkalimetalldichalkogenide β-Na_2S_2, K_2S_2, α-Rb_2S_2, β-Rb_2S_2, K_2Se_2, Rb_2Se_2, α-K_2Te_2, β-K_2Te_2 und Rb_2Te_2*, *Z. fur Anorg. Allg. Chem.* **619** (1993), no. 3 476–488.

[134] V. Franco, J. Blázquez, B. Ingale, and A. Conde, *The magnetocaloric effect and magnetic refrigeration near room temperature: materials and models*, *Annual Review of Materials Research* **42** (2012) 305–342.

[135] O. Gutfleisch, M. A. Willard, E. Brück, C. H. Chen, S. Sankar, and J. P. Liu, *Magnetic materials and devices for the 21st century: stronger, lighter, and more energy efficient*, *Advanced materials* **23** (2011), no. 7 821–842.

[136] M. Balli, S. Jandl, P. Fournier, and A. Kedous-Lebouc, *Advanced materials for magnetic cooling: Fundamentals and practical aspects*, *Appl. Phys. Rev.* **4** (2017), no. 2 021305.

[137] J. Lyubina, *Magnetocaloric materials for energy efficient cooling*, *Journal of Physics D: Applied Physics* **50** (2017), no. 5 053002.

[138] P. J. Shirron, *Cooling capabilities of adiabatic demagnetization refrigerators*, *Journal of Low Temperature Physics* **148** (2007) 915–920.

[139] P. J. Shirron, M. O. Kimball, D. J. Fixsen, A. J. Kogut, X. Li, and M. J. DiPirro, *Design of the pixie adiabatic demagnetization refrigerators*, *Cryogenics* **52** (2012), no. 4-6 140–144.

[140] T. J. Sato, D. Okuyama, and H. Kimura, *Tiny adiabatic-demagnetization refrigerator for a commercial superconducting quantum interference device magnetometer*, *Review of Scientific Instruments* **87** (2016), no. 12.

[141] C. Aprea, A. Greco, A. Maiorino, and C. Masselli, *Magnetic refrigeration: an eco-friendly technology for the refrigeration at room temperature*, in *Journal of Physics: Conference Series*, vol. 655, p. 012026, IOP Publishing, 2015.

[142] R. Teyber, J. Holladay, K. Meinhardt, E. Polikarpov, E. Thomsen, J. Cui, A. Rowe, and J. Barclay, *Performance investigation of a high-field active magnetic regenerator*, *Applied Energy* **236** (2019) 426–436.

[143] S. M. Safdarnejad, J. D. Hedengren, and L. L. Baxter, *Plant-level dynamic optimization of cryogenic carbon capture with conventional and renewable power sources*, *Applied Energy* **149** (2015) 354–366.

[144] L. Patra and B. Liao, *Indirect exchange interaction leads to large lattice contribution to magnetocaloric entropy change*, *Physical Review Letters* **131** (2023) 066703.

[145] E. A. Olivetti, J. M. Cole, E. Kim, O. Kononova, G. Ceder, T. Y.-J. Han, and A. M. Hiszpanski, *Data-driven materials research enabled by natural language processing and information extraction*, *Applied Physics Reviews* **7** (2020), no. 4.

[146] V. Venugopal, S. Sahoo, M. Zaki, M. Agarwal, N. N. Gosvami, and N. A. Krishnan, *Looking through glass: Knowledge discovery from materials science literature using natural language processing*, *Patterns* **2** (2021), no. 7.

[147] C. J. Court and J. M. Cole, *Auto-generated materials database of Curie and Néel temperatures via semi-supervised relationship extraction*, *Scientific Data* **5** (2018), no. 1 1–12.

[148] B. Min, H. Ross, E. Sulem, A. P. B. Veyseh, T. H. Nguyen, O. Sainz, E. Agirre, I. Heintz, and D. Roth, *Recent advances in natural language processing via large pre-trained language models: A survey*, *ACM Computing Surveys* **56** (2023), no. 2 1–40.

[149] M. P. Polak and D. Morgan, *Extracting accurate materials data from research papers with conversational language models and prompt engineering*, *Nature Communications* **15** (2024), no. 1 1569.

[150] V. P. Antropov, M. Katsnelson, M. Van Schilfgaarde, and B. Harmon, *Ab initio spin dynamics in magnets*, *Phys. Rev. Lett* **75** (1995), no. 4 729.

[151] B. Skubic, J. Hellsvik, L. Nordström, and O. Eriksson, *A method for atomistic spin dynamics simulations: implementation and examples*, *J. Phys. Cond. Matt.* **20** (2008), no. 31 315203.

[152] J. Dagdelen, A. Dunn, S. Lee, N. Walker, A. S. Rosen, G. Ceder, K. A. Persson, and A. Jain, *Structured information extraction from scientific text with large language models*, *Nature Communications* **15** (2024), no. 1 1418.

[153] R. L. Logan IV, I. Balažević, E. Wallace, F. Petroni, S. Singh, and S. Riedel, *Cutting down on prompts and parameters: Simple few-shot learning with language models*, *arXiv preprint arXiv:2106.13353* (2021).

[154] G. Roemer, A. Li, U. Mahmood, L. Dauer, and M. Bellamy, *Artificial intelligence model GPT4 narrowly fails simulated radiological protection exam*, *Journal of Radiological Protection* **44** (2024), no. 1 013502.

[155] M. Lee, *A Mathematical Investigation of Hallucination and Creativity in GPT Models*, *Mathematics* **11** (2023), no. 10.

[156] S. Hcini, A. Dhahri, S. Zemni, and M. L. Bouazizi, *Structural, magnetic and magnetocaloric properties and analysis of MCE using the mean-field theory of Mg–Co ferrite with Ni substitution*, *Journal of Materials Science: Materials in Electronics* **30** (2019) 6127–6138.

[157] C. Sammut and G. I. Webb, *Encyclopedia of machine learning*. Springer Science & Business Media, 2011.

[158] G. Kresse and J. Furthmüller, *Efficient iterative schemes for ab-initio total-energy calculations using a plane-wave basis set*, *Phys. Rev. B* **54** (1996), no. 16 11169.

[159] H. J. Monkhorst and J. D. Pack, *Special points for Brillouin-zone integrations*, *Phys. Rev. B* **13** (June, 1976) 5188–5192.

[160] C.-G. Duan, R. Sabirianov, W.-N. Mei, P. A. Dowben, S. Jaswal, and E. Y. Tsymbal, *Electronic, magnetic and transport properties of rare-earth monopnictides*, *Journal of Physics: Condensed Matter* **19** (2007), no. 31 315220.

[161] X. Wang, G. Ding, Z. Cheng, H. Yuan, X.-L. Wang, T. Yang, R. Khenata, and W. Wang, *$R\bar{3}c$-type $LnNiO_3$ (ln= La, Ce, Nd, Pm, Gd, Tb, Dy, Ho, Er, Lu) half-metals with multiple Dirac cones: A potential class of advanced spintronic materials*, *IUCrJ* **6** (2019), no. 6 990–995.

[162] R. F. Evans, W. J. Fan, P. Chureemart, T. A. Ostler, M. O. Ellis, and R. W. Chantrell, *Atomistic spin model simulations of magnetic nanomaterials*, *J. Phys. Condens. Matter* **26** (2014), no. 10 103202.

[163] J. L. García-Palacios and F. J. Lázaro, *Langevin-dynamics study of the dynamical properties of small magnetic particles*, *Physical Review B* **58** (1998), no. 22 14937.

[164] A. Togo, L. Chaput, T. Tadano, and I. Tanaka, *Implementation strategies in phonopy and phono3py*, *J. Phys. Condens. Matter* **35** (2023), no. 35 353001.

[165] A. Togo, *First-principles phonon calculations with phonopy and phono3py*, *J. Phys. Soc. Jpn.* **92** (2023), no. 1 012001.

[166] G. Wang, *Magnetic and calorimetric study of the magnetocaloric effect in intermetallics exhibiting first-order magnetostructural transitions*, vol. 99. Universidad de Zaragoza, 2012.

[167] B. Alho, P. Ribeiro, R. de Oliveira, V. de Sousa, E. Nóbrega, B. Margato, J. da Silva, and P. von Ranke, *Mean-field modeling of magnetocaloric effect of antiferromagnetic compounds*, *IEEE Magnetics Letters* **13** (2022) 1–4.

[168] P. Von Ranke, N. De Oliveira, B. Alho, E. Plaza, V. De Sousa, L. Caron, and M. Reis, *Understanding the inverse magnetocaloric effect in antiferro-and ferrimagnetic arrangements*, *Journal of Physics: Condensed Matter* **21** (2009), no. 5 056004.

[169] E. Z. Valiev, *Entropy and magnetocaloric effect in ferromagnets and antiferromagnets*, *The Physics of Metals and Metallography* **104** (2007) 8–12.

[170] S. Süllow, I. Prasad, M. Aronson, J. Sarrao, Z. Fisk, D. Hristova, A. Lacerda, M. Hundley, A. Vigliante, and D. Gibbs, *Structure and magnetic order of EuB_6*, *Physical Review B* **57** (1998), no. 10 5860.

[171] M. Danebrock, W. Jeitschko, A. Witte, and R. Pöttgen, *Magnetic properties of the carbides RCoC (R= Y, Gd-Er, Lu) with YCoC-type structure*, *Journal of Physics and Chemistry of Solids* **56** (1995), no. 6 807–811.

[172] A. Castets, D. Gignoux, and J. Gomez-Sal, *Magnetic properties and structures of equiatomic rare earth-platinum compounds RPt (R= Gd, Tb, Dy, Ho, Er, Tm)*, *Journal of Solid State Chemistry* **31** (1980), no. 2 197–207.

[173] R. Welter, A. Morozkin, and K. Halich, *A magnetic study of the CeScSi-type RZrSb compounds (R= Gd–Tm). Large coercive field at low temperature in TbZrSb and DyZrSb*, *Journal of magnetism and magnetic materials* **257** (2003), no. 1 44–50.

[174] D. Ryan, J. Cadogan, S. Xu, Z. Xu, and G. Cao, *Magnetic structure of $EuFe_2P_2$ studied by neutron powder diffraction*, *Physical Review B* **83** (2011), no. 13 132403.

[175] J. Greedan, *The magnetic properties of $EuLiH_3$*, *Journal of Physics and Chemistry of Solids* **32** (1971), no. 4 819–823.

[176] L. Petit, Z. Szotek, J. Jackson, M. Lüders, D. Paudyal, Y. Mudryk, V. Pecharsky, K. Gschneidner Jr, and J. Staunton, *Magnetic properties of Gd intermetallics*, *Journal of Magnetism and Magnetic Materials* **448** (2018) 9–12.

[177] T. A. Yamamoto, T. Nakagawa, K. Sako, T. Arakawa, and H. Nitani, *Magnetocaloric effect of rare earth mono-nitrides, TbN and HoN*, *Journal of alloys and compounds* **376** (2004), no. 1-2 17–22.

[178] P. von Ranke, T. Alvarenga, E. Nóbrega, B. Alho, P. Ribeiro, A. M. G. Carvalho, V. de Sousa, A. Caldas, and N. de Oliveira, *Investigation on the magnetocaloric effect in TbN compound*, *Journal of magnetism and magnetic materials* **341** (2013) 138–141.