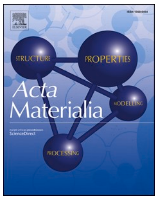
Acta Materialia 281 (2024) 120342

Contents lists available at ScienceDirect

Acta Materialia

journal homepage: www.elsevier.com/locate/actamat



Full length article

Process optimization on kesterite-based ceramics for enhancing their thermoelectric performances assisted by active machine learning approach: A tool for metal-sulfide ceramics development

Cédric Bourgès [a,*], Guillaume Lambard [b,*], Naoki Sato [c], Makoto Tachibana [c], Satoshi Ishii [c,d], Takao Mori [c,d]

[a] International Center for Young Scientists (ICYS), National Institute for Materials Science (NIMS), 1-1 Namiki, Tsukuba 305-0044, Japan
[b] Center for Basic Research on Materials (CBRM), National Institute for Materials Science (NIMS), Namiki 1-1, Tsukuba 305-0044, Ibaraki, Japan
[c] Research Center for Materials Nanoarchitectonics (MANA), National Institute for Materials Science (NIMS), 1-1 Namiki, Tsukuba 305–0044, Ibaraki, Japan
[d] Graduate School of Pure and Applied Sciences, University of Tsukuba, 1-1-1 Tennoudai, Tsukuba 305-8577, Japan

ARTICLE INFO

Keywords:
Kesterite
Machine learning
Process
Thermoelectric
Ceramic

ABSTRACT

The thermal process parameters are crucial in metal-sulfides ceramics as they affect significantly the resulting physico-chemical properties. In the present work, we investigated the sintering effect in the kesterite $Cu_{2.125}Zn_{0.875}SnS_4$ on its structural, microstructural, and thermoelectric (TE) properties to highlight the non-negligible contribution of the thermal process often ignored in metal-sulfide ceramics. For this purpose, we developed an approach combining data science with the conventional material experiment/theory approach which can be used as a tool to shortcut the time-consuming steps of TE material optimization. We confirmed that the optimization and control of the densification process is critical in unravelling the highest potential on metal sulfide TE ceramics with a non-negligible increase of its zT up to 60 %. We propose a scientific tool, the synergic combination of active machine learning with conventional chemistry/theory approaches, to either identify the most proficient sintering process as well as the process to avoid the degradation of the metal-sulfide ceramic properties and thus in a shorten number of experiments. This approach can be extended not only to other metal-sulfide ceramics for thermoelectricity but also to other research fields.

1. Introduction

In the present 21st century, human development faces multiple major challenges that need to be overcome to complete the construction of society 5.0 [1]. Within those challenges, the development of energy saving, or alternative sustainable energy source technology is at the center of an extensive research area.

Thermoelectric (TE) power generation constitutes one of the attractive technological solutions which can reduce energy loss by partially recovering the massive waste heat emitted by the current society (transport, industry, housing) and converting it to electrical power for contributing to a sustainable future [2,3]. The dimensionless figure of merit $zT = PFT\kappa^{-1} = S^2\sigma T\kappa^{-1}$ (PF, power factor; T, absolute temperature; S, Seebeck coefficient; σ, electrical conductivity; κ, thermal conductivity) can conveniently quantify the TE conversion efficiency of a material. Numerous materials have been reported with zT values around or above the unity in the intermediate temperature range (300 – 800 K). Among them, we can find Bi_2Te_3 [4,5], a constituent of most of the current commercial TE modules, but also the PbTe [6,7], and Mg_3Sb_2-based compounds [8,9]. However, the main drawbacks limiting their development at an industrial scale lie in the costly, scarce, and/or toxic elements which compose those materials. As alternative materials, the oxide and sulfide compounds have attracted some interest in the field due to their cheap and/or eco-friendlier compositions suitable for the sustainability requirement [10]. Among these families, the oxides are represented by ZnO [11,12], or calcium cobaltite $Ca_3Co_4O_9$ [13,14] that are attractive for their high stability in the high-temperature range (T > 800 K) but which still have weak performance in the mid-temperature (300 – 800 K). Sulfide based ceramics are attractive in the regard. Indeed, a few natural Cu-based sulfides minerals exhibit outstanding performance such as the binary $Cu_{2\text{-}x}S$ [15–17], the tetrahedrite $Cu_{12\text{-}x}Tr_xSb_4S_{13}$ (Tr = Mn, Fe, Co, Ni, Zn) [18–20] and also the

---

* Corresponding authors.
E-mail addresses: BOURGES.Cedric@nims.go.jp (C. Bourgès), LAMBARD.Guillaume@nims.go.jp (G. Lambard).

https://doi.org/10.1016/j.actamat.2024.120342
Received 2 November 2023; Received in revised form 1 August 2024; Accepted 26 August 2024
Available online 29 August 2024

recent colusites $Cu_{26}T_2M_6S_{32}$ ($T = V$, Nb, Ta, Cr, Mo, W, Ti, $M =$ Ge, Sn, Sb) [21–24], which all reached or exceeded the zT value of 1. This drove, during the last decade, extensive investigation on other multi-cationic Cu-based sulfides such as bornites Cu_5FeS_4 [25,26], chalcopyrites $CuFeS_2$ [27–30], Cu-based thiospinels [31–34], or kesterite-stannite type and derivates Cu_2TrMsS_4 (Tr = Fe, Co, Ni, Zn, Ms = Ge, Sn) [35–40].

The kesterite type $Cu_2ZnSn(S,Se)_4$ is already known and reported as one of the attractive new generations of high-performance photovoltaic materials for solar cell applications, thanks to its appropriate bandgap (1–1.5 eV) with high light absorption coefficients and thermodynamically stable adamantine structure [41]. However, it has been poorly investigated for TE applications due to its intrinsic semi-conducting character. Interestingly, kesterite shows a phase transition around $T \approx 500$ K from ordered to a cationic disordered structure (**cf. 3.3.**) [42]. The phase transition has a significant effect on the TE properties of the material unravelling its potential for the field, especially through the realization of a $zT = 1.6$ at 800 K on a single crystal sample. [36,40,43] Nevertheless, polycrystalline samples represent a significant challenge due to the high sensitivity of the synthesis and densification process susceptible to influence the final bulk kesterite structure and properties and explain the discrepancies in the current reports as for example observed between the report of Isotta *et al.* and the Jiang *et al.* which show a different trend in their respective TE properties [35,39,43]. It appears therefore that the process control can be a key factor obtaining of the ideal single crystal performance. Unfortunately, such studies are rarely engaged and reported as it is time and resource consuming. For such purposes, the field of machine learning provides novel tools suitable to perform process optimization [44,45]. On the first hand, it can reduce the number of experiments required to find the ideal set of process parameters that improve TE properties and, on the second hand, propose statistical relationships between the process parameters and the targeted physical properties. Notably, the latter advantage supports the establishment of dependencies which could appear pertinent to the understanding of a physicochemical system like in the present study. To this end, the active learning pipeline is a general framework that comprises: a “learning” step based on past experiments and during which a statistical modelization of a system {process parameters, targeted properties} is attempted by using experimental data to train a machine learning model that capture relationships between parameters and properties; And an “active” step in which a set of experimental actions, derived from the learned model of the system, are performed and are supposed to bring the system closer to an objective. Bayesian optimization leverages this model-building to guide the choice of parameter sets. Multiple learning-acting cycles constitute an active learning pipeline. A trial-and-error approach can be viewed as following an active learning scheme if, at each cycle, it includes both an update of theoretical/empirical knowledge and informed forward actions that modify the experimental design to meet an objective. However, systematically, and cyclically updating knowledge and actions can be challenging, especially when the objective has a nonlinear relationship with many experimental parameters. In this sense, an active learning pipeline assisted by machine learning and Bayesian optimization (ALMLBO) pipeline offers the possibility to optimize salient input features among 5 experimentally tunable process parameters for achieving high zT in kesterite with composition $Cu_{2.125}Zn_{0.875}SnS_4$. Here, we report a high $zT = 0.44$ at 775 K with high purity precursors after just 4 cycles of ALMLBO. It is crucial to emphasize that ALMLBO is here not intended as a standalone predictor, but rather as a guide for experimental design. This approach efficiently explores the vast parameter space of sintering conditions, identifying promising directions for experimentation that might be overlooked by traditional methods. This is particularly valuable in complex material systems where the relationships between processing parameters and properties are not always straightforward.

Herein, we focused on the investigation of process influence on the kesterite, commonly named CZTS, with formulation $Cu_{2.125}Zn_{0.875}SnS_4$ by using an ALMLBO approach [46] to obtain enhanced TE properties and highlight an ideal set of process parameters for this compound. Based on preliminary investigation and literature screening, the composition $Cu_{2.125}Zn_{0.875}SnS_4$ has been selected to obtain initial reasonable TE properties, especially electronic transport properties. To limit the number of free parameters and possibly learn on the influence of the densification process through the lens of ALMLBO, the synthesis method has been limited to the mechanical alloying (MA) process. Indeed, the MA synthesis is commonly used to produce sulfide powder thanks to its intrinsic advantages. In comparison to the commercial powder and/or conventional sealed tube synthesis, the crystallinity of the MA powder is usually reduced, though the composition is more stoichiometric and homogeneous with lower grain size and narrower particle distribution (< 1 μm). These features can allow decreasing the sintering temperature/sintering time to have better control of the possible induced off-stoichiometry which can occur and therefore relate more accurately to the influence of the densification. Besides, the MA is attractive due to the high reproducibility and the possibility of large batch production particularly adapted for process investigation. The samples have been structurally and microstructurally characterized as well as their TE properties from 300 to 775 K. The experimental results were rationalized with *ab initio* calculations to sustain the experimental observation/interpretation.

2. Methods

2.1. Samples preparation

The samples of $Cu_{2.125}Zn_{0.875}SnS_4$ were prepared by mechanical alloying process before being consolidated by Spark Plasma Sintering (SPS). The raw powders were synthesized by mixing 5 g of the stoichiometric ratio of the elemental Copper (3Nup, powder, 75 μm pass), Zinc (99.9 %, powder, 75 – 150 μm), Tin (99.8 %, powder, < 45 μm) and Sulfur (≥ 99.99 %, Flake) in a SPEX stainless steel grinding vial filled with two 10 mmØ and four 5 mmØ stainless steel ball during 3 h in a SPEX 8000D mixer/mill. The resulting powder were sintered (SPS-1080 – SPS Syntex Inc.) in a 10 mm diameter graphite dies and a partial vacuum atmosphere following variable sintering parameters according to the ALMLBO optimization (cf. ALMLBO part). The archetype sintering cycles presented in the article (**cf. 3.2.**) correspond to the following parameters $\boldsymbol{\Theta} = \{HR, SintT, Stept, CR, UP\}$: Reference (*HR*: 50 K/min, *SintT*: 773 K, *Stept*: 15 min; *CR*: 50 K/min, *UP*: 50 MPa); Degrading Process: (*HR*: 25 K/min, *SintT*: 1073 K, *Stept*: 45 min; *CR*: 75 K/min, *UP*: 40 MPa), and MLProcess (*HR*: 150 K/min, *SintT*: 673 K, *Stept*: 45 min; *CR*: 75 K/min, *UP*: 40 MPa). The sintered pellets were then polished and cut to the required shapes and dimensions for the property’s measurements. The optimized sample after ALMLBO were reproduced with high purity Copper (3Nup, powder, 75 μm pass), Zinc / Sulfur (4 N, powder, 75 μm pass) and Sn (4 N, powder, 63 μm pass) from Kojundo company and new SPEX hardened steel grinding vial with same ball contents than previously described to probe the possible influence of chemical’s purity and jar contamination (Sample MLP+HPC). The samples were reproduced on a second SPS (Dr. Sinter Lab Jr. series 322Lx, Fuji Electronic Industrial Co., LTD.) to probe the reproducibility. To clarify the text of the manuscript and ensure ease of understanding, we have labeled the samples in both the manuscript and the figures as follows: the reference sample is labeled as "Reference"; the sample related to the machine learning process is labeled as "MLProcess"; the sample associated with the machine learning process that uses a high purity chemical is labeled as "MLProcess+HPC"; and the sample undergoing a degrading process is labeled as "DP."

2.2. Characterizations

The phase purity and crystal structure were characterized by X-ray diffraction (XRD) with the prescriptive θ–2θ Bragg-Brentano diffractometer over an angle of 10°–140° (Smart Lab3 Rigaku Corporation) with a step size of 0.02° and a step time of 2°/min. X-ray powder diffraction patterns were refined by Rietveld analysis using the FullProf and WinPLOTR software packages [47,48]. The shape of the diffraction peaks was modeled using a Thompson–Cox–Hastings pseudo-Voigt profile function [49]. Zero-point shift, asymmetry parameters and lattice parameters were systematically refined, and the background contribution was manually estimated and corrected.

Observations of a fractured cross-section of densified pellets by scanning electron microscopy (SEM) were performed using a Hitachi SU-8230 and the component analysis was proceeded by energy dispersive spectroscopy (EDS) on mirror polished surface with a Hitachi TM3000 Tabletop SEM.

The carrier concentration and mobility were obtained from Hall Effect measurements in AC transport by using a PPMS (Quantum design) from −5 to +5 T.

The electrical conductivity (σ) and Seebeck coefficient (S) were simultaneously measured by the four-probe method from 300 K up to 725 K using a ZEM-2 (ULVAC Advance-Riko) device under partial helium pressure.

The temperature-dependent thermal diffusivity, α, was measured using the laser flash diffusivity method in LFA-467 Hyperflash (Netzsch) on graphite coated disk-shaped samples of 10 mm diameter and ~2 mm thickness. The temperature-dependent heat capacity, C_p, was determined by DSC methods using a sapphire standard reference. The total thermal conductivity was calculated using the formula, $\kappa = \alpha \times C_p \times \rho$, where ρ is the density of the sample (measured using Archimedes' principle). In some cases, to better understand the thermal transport properties, the contributions from electronic and lattice parts were calculated. The lattice thermal conductivity (κ_{latt}) was estimated from κ by subtracting the electronic contribution (κ_e) *via* the Wiedemann-Franz law, as in the following equation,

$$\kappa_e = L\sigma T \tag{1}$$

where κ_e is the electronic thermal conductivity, and L is the Lorenz number.

The uncertainty in the results for the values of the Seebeck is 6 %, electrical resistivity 8 %, power factor 11.2 %, thermal conductivity 10 %, and for the overall zT was 17.5 % according to a reference round robin test [50].

2.3. Computational method

The DFT calculations were performed using Quantum ESPRESSO (QE) [51,52] package with projector augmented wave pseudopotentials [53]. The generalized gradient approximation functional with Perdew-Burke-Ernzerhof parametrization (GGA-PBE) [54] was chosen for the exchange and correlation potentials. We applied the Hubbard U correction to the strongly correlated Cu 3d, Zn 3d, and Sn 4d orbitals. We chose U_{eff} ($= U - V$) values as 5.2 eV for Cu, and 6.5 eV for Zn, as it reproduces the experimental lattice accurately (Table S1). Spin-orbit coupling was not considered in the present study. The conventional unit cell with tetragonal lattice containing 16 atoms was fully relaxed until the residual forces became less than 10^{-6} eV $Å^{-1}$. The cutoff energy of 130 Ry for the plane wave basis and 8 × 8 × 4 k-grid were used for the structure relaxation.

To extract the second and third order interatomic force constants (IFCs), we used finite difference method with 4 × 4 × 2 and 2 × 2 × 1 supercells based on the fully relaxed conventional cell to create displacement-force datasets. The magnitude of atomic displacements was set at 0.01 Å and 0.04 Å for calculating the second and third order IFCs, respectively. For the third order IFCs calculation, we chose the cutoff radius as 5.44 Å for choosing triplets within which interaction is considered. The IFCs were obtained using ALAMODE [55]. The phonon dispersion relation and DOS were calculated from the second order IFC (**cf. 3.4.**). Dielectric constant and Born effective charge were calculated in the perturbation framework implemented in the QE package for correcting phonon frequencies considering long-range interaction. ALAMODE code is also used to calculate the lattice thermal conductivity by solving the phonon Boltzmann transport equation within the single mode relaxation time approximation. The sampling grid of 18 × 18 × 18 was applied after confirming convergence of the calculated lattice thermal conductivity with respect to the grid density.

3. Results and discussion

3.1. Active learning pipeline assisted by machine learning and Bayesian optimization (ALMLBO)

First, a synthesis protocol is established (cf. experimental part) for building an initial dataset composed of the measurable TE properties $\mathbf{\Pi} = \{S, \rho\ (=1/\sigma), \kappa_{tot}\}$, summarized to the $zT = S^2\rho^{-1}T\kappa_{tot}{}^{-1}$, and tunable sintering process parameters $\mathbf{\Theta} = \{HR, SintT, Stept, CR, UP\}$ of kesterite samples, where: *HR* and *CR* are the heating and cooling rates (K/min), respectively; *SintT* and *Stept* are the sintering temperature (K) and step time (min), respectively; *UP* is the uniaxial pressure (MPa).

The ALMLBO pipeline was kickstarted with only $n = 11$ samples $\{\mathbf{\Theta}_i, \mathbf{\Pi}_i \mid i = 1,\ldots,11\}$, initially selected on a trial-and-error basis to obtain single-phase kesterite samples with satisfactory TE properties $\mathbf{\Pi}$, each $\mathbf{\Pi}_i$ being identified as a set of average values $\{S, \rho, \kappa_{tot}\}$ measured from the same parent sample of kesterite. Each tunable Spark Plasma Sintering (SPS) process parameter $\theta_i \in \mathbf{\Theta}_i$, issued from all possible $\mathbf{\Theta}$, was restricted to vary within a domain d_i within a resolution δd_i, both chosen to be experimentally feasible, as summarized in Table 1. At this stage, $N = 36{,}608$ combinations of $\mathbf{\Theta}i$ were experimentally possible such that $N = \prod_i \{[(max\_d_i - min\_d_i) / \delta d_i]+1\}$, where $\prod_i$ is a product across i tunable process parameters, max_d_i and min_d_i are the maximum and minimum values, respectively, of a domain d_i.

Then, the ALMLBO pipeline serves as a bridge between theory and experiment, enabling rapid iteration and knowledge discovery. By suggesting optimal parameters for each experimental cycle, it accelerates the optimization process and allows for a more thorough exploration of the material's potential within practical constraints. A cycle of the ALMLBO pipeline, summarized in Fig. 1**a)**, is conducted as follows: i) a set $\boldsymbol{\varepsilon}_{RF}$ of an ensemble of n_{RF} Random Forests (RF) regressors [56], with $n_{RF} = n$ available experimental samples and default hyper-parameters of the RF models assigned as in Pedregosa et al. [58], are trained to predict zT_i derived from measured $\mathbf{\Pi}i$, following a leave-one-out cross-validation (LOO—CV) scheme [57], and using a mean-squared-error as the loss function to be minimized during training. RF regressors have been preferably chosen among other machine learning models, like ElasticNet, XGBoost, and LightGBM, for their performance following a preliminary study. Overall, a reasonable average root-mean-square-error (RMSE) ~ 0.04 ± 0.03 on the zT is achieved after the 4th cycle of ALMLBO. A visualization of the evolution of the RF models performance following a LOO—CV scheme can be found in **Fig. S1** of the supplementary materials. Overall, the prediction

Table 1
List of tunable process parameters $\theta_i \in \mathbf{\Theta}_i$, issued from $\mathbf{\Theta}$ = {HR, SintT, Stept, CR, UP}, with their accessible domain d_i and resolution δd_i chosen for the densification phase via SPS of the kesterite $Cu_{2.125}Zn_{0.875}SnS_4$. HR and CR are the heating and cooling rates (K/min), respectively; SintT and Stept are the sintering temperature (K) and step time (min), respectively; UP is the uniaxial pressure (MPa).

Process parameter θ_i	Domain d_i	Resolution δd_i
HR	[25, 200] K/min	25 K/min
SintT	[573, 1073] K	50 K
Stept	[0, 60] min	5 min
CR	[25, 200] K/min	25 K/min
UP	[40, 80] MPa	10 MPa

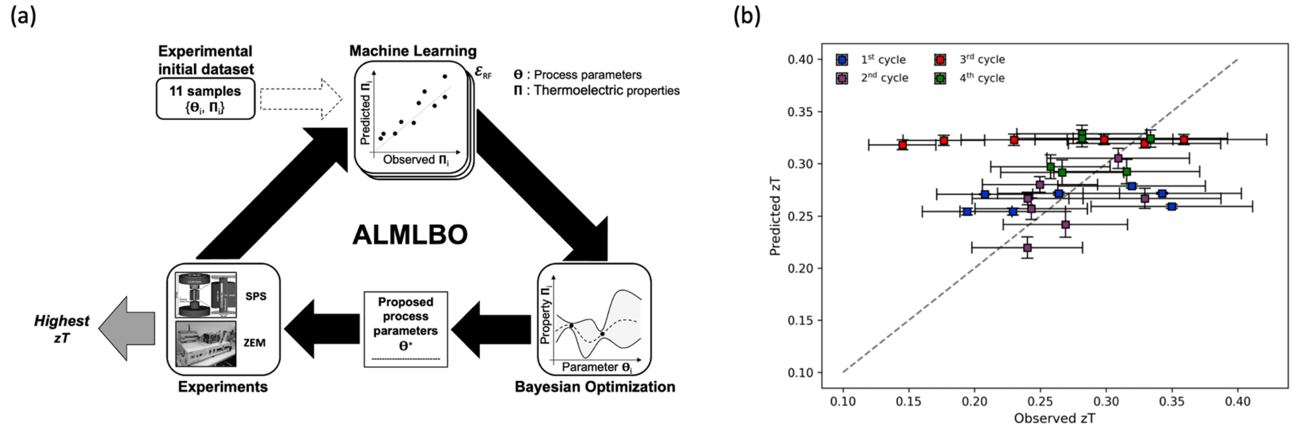


Fig. 1. a) Active learning assisted by machine learning and Bayesian optimization (ALMLBO) applied to the optimization of SPS process parameters of the kesterite $Cu_{2.125}Zn_{0.875}SnS_4$ for achieving high *zT*; b) Comparison between predicted (Random Forests regression [56]) and average *zT* observed after a 1st, 2nd, 3rd, and 4th ALMLBO cycle (blue, purple, red, and green squares, respectively) with vertical and horizontal error bars representing a standard deviation related to a leave-one-out [57] cross-validation, and computed from $m = 1$–2 experimental measurements and the systematic uncertainty of ~17.5 % [50], respectively (the dashed line shows a perfect prediction for the purpose of visual guidance only).

performance on *zT* are mainly degraded by the poorer prediction on ρ and κ_{tot}. In comparison, the prediction on *S* are more accurate revealing closer non-linear relations to the SPS process parameter than ρ and κ_{tot} for which predictions models remain to be improved. It is worthy to notice that ρ and κ_{tot} are parameters which are not only intrinsic to the material crystal structure but also affect by extrinsic parameter such as the density or the microstructure of the ceramic compared to the *S*. Its can be one of the explanations regarding the difficulty of prediction which can be improved. For a detailed visualization of the individual prediction performance on all the properties **Π, Fig. S1-S4** in the supplementary materials can be examined. Here, all parameters in **Θ** are kept for optimization purposes, and the remaining parameters $\mathbf{\Theta}' \not\subset \mathbf{\Theta}$, i.e., related to powder mixing and ball milling, are tuned to their most efficient values following our know-how of the present experimental setting; ii) A set $\{\mathbf{\Theta}, \mathbf{\Pi}\}|_r$ of 1000 samples, with $\mathbf{\Theta}_i \in d_i$ (see Table 1) for the ith sample, is randomly evaluated through $\boldsymbol{\varepsilon}_{RF}$ to initialize a following Gaussian process regressor at step iii), and relax from any experimental biases and/or local optima that would have occurred during the preparation of the dataset issued from a previous cycle; iii) A Gaussian process is trained on $\{\mathbf{\Theta}, \mathbf{\Pi}\}|_r$ that serves as a surrogate function f_s to exploit and explore the space of **Θ** parameters; (iv) A set $\mathbf{\Theta}^*$ of 50 samples that maximizes the expected improvement [56], chosen as acquisition function, is proposed for experimental feedback. Steps ii)-iii) represent the Bayesian optimization (BO) part in the ALMLBO pipeline here; iv) A set $\mathbf{\Theta}^*$ of 6–7 samples are chosen for experimental feasibility, are experimentally evaluated, and corresponding $\mathbf{\Pi}^*$ properties are reported. If $\mathbf{\Pi}^*$ are judged high enough, and $\mathbf{\Theta}^*$ become stable after an nth ALMLBO cycle, the ALMLBO pipeline is stopped. Otherwise, another cycle i)-iv) is started by adding newly acquired experimental data to the whole dataset. Fig. 1b) shows the evolution of predicted *zT* values over 4 ALMLBO cycles (blue, purple, red, and green squares, respectively) in comparison to their experimentally observed counterparts for evaluated samples $\mathbf{\Theta}^*$ following the BO step. Indeed, Fig. 1b.) illustrates the prediction performance of the set $\boldsymbol{\varepsilon}_{RF}$ of trained RF models on out-of-distribution data, i.e., with yet unseen sets of SPS process parameters **Θ**. It can be observed that the ALMLBO pipeline passes by successive steps of exploration and exploitation. The ensemble of predictors $\boldsymbol{\varepsilon}_{RF}$ occurs to be in a saturation regime for the 1st and 3rd ALMLBO cycles, a regime during which predictions on *zT* are mostly high but equals, illustrating the impossibility for RF models to predict zT values from unseen **Θ** higher than already measured *zT* present in the dataset used for training, checking the boundaries of yet unexplored **Θ** space. Then, $\boldsymbol{\varepsilon}_{RF}$ gets into an exploitation regime in the 2nd and 4th ALMLBO cycles, where predicted *zT* values are statistically consistent with measured counterparts, as mainly illustrated by the cover of experimental uncertainties.

Fig. 2**a)** shows the evolution of the **Θ** features importance [57] in predicting *zT* with $\boldsymbol{\varepsilon}_{RF}$ across the different cycles of ALMLBO, i.e., from the initial dataset (Init.) to a dataset containing all the data acquired after 4 ALMLBO cycles (Init. w. 1–4 cycles). Overall, the SinT quickly occurs to be a salient parameter influencing *zT* with a quick decrease in UP importance from the 1st ALMLBO cycle. This last statement on UP importance illustrates the appearance of spurious statistical correlations that may occur when other influencing **Θ** parameters are initially fixed, like Stept, or a regime where *zT* barely varies, like with CR, or where the effect of two **Θ** features cancels out each other, like HR and SinT, in the initial data. Nevertheless, apparent spurious statistical correlations quickly vanish from the 1st ALMLBO where the SPS process parameter SinT becomes dominant, the following ALMLBO cycles confirming the importance of SinT in primarily affecting *zT*. This **Θ** feature importance is confirmed across the whole set of **Π** properties as illustrated in **Fig. S5** in the supplementary materials, the SPS process parameters {Stept, HR, UP, CR}, here listed in the decreasing order of their average importance on *zT* predictions, sharing the remaining individual effects on *zT*. To even further validate the importance of the **Θ** features and ensure they are truly informative rather than relying on spurious correlations, we conducted additional permutation importance tests [60,61]. This technique randomly shuffles each feature and measures the decrease in the RF models' performance, providing a robust measure of feature importance that is independent of the RF models' internal mechanism. **Fig. S5-bis** in the supplementary materials shows the permutation importance for each feature. Notably, the consistency between these permutation importances and feature importances in **Fig. S5** further validates the robustness of our consequent feature interpretation above and the reliability of our ALMLBO's decision-making process. Fig. 2**b)** illustrates the partial dependence [59] of *zT* on SinT from 550 to 1100 K. A SinT = 673 K, where $zT = 0.36@675$ K (0.40 with high purity chemicals, see below) is maximised during the 3rd ALMLBO, shows to be optimal in returning high, relatively stable *zT* values for the kesterite $Cu_{2.125}Zn_{0.875}SnS_4$ as illustrated by a narrower shaded area surrounding the mean (median) value. Values of SinT > 1023 K would also permit an increase in *zT* with, nevertheless a much higher uncertainty. For the remaining SPS process parameters {Stept, HR, UP, CR}, **Fig. S6** illustrates the partial dependence of *zT* as well showing that CR and UP barely affect *zT* in comparison to SinT, as well as Stept except for Stept > 40 mins where a Stept = 45 mins is used for the reported optimal $zT = 0.36@675$ K, and for HR that tends to increase *zT* if increased with a constant stability along the imposed regime.

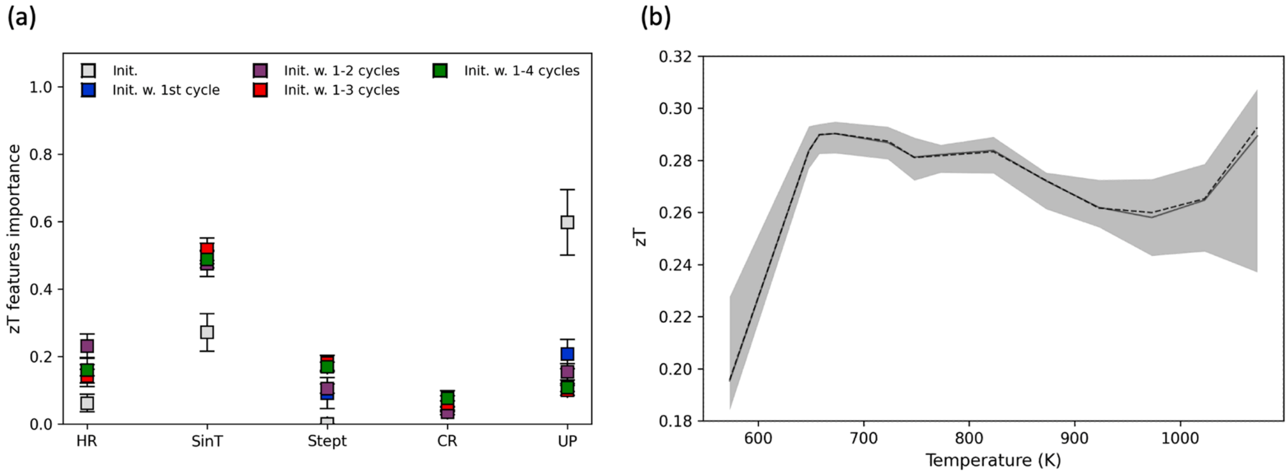


Fig. 2. a) Distribution of average features importance [57] for the SPS process parameters θ regarding the prediction of the figure-of-merit, zT, based on the pre-ALMLBO initial data (gray squares) extended with data from the 1st, 2nd, 3rd, and 4th ALMLBO cycles (blue, purple, red, and green squares, respectively). Mean values and error bars are provided according to the leave-one-out [57] cross-validation used for Random Forests [56] models training. HR and CR are the heating and cooling rates (K/min), respectively; SinT and Stept are the sintering temperature (K) and step time (min), respectively; UP is the uniaxial pressure (MPa); b) Partial dependence [59] of the figure-of-merit, zT, on the sintering temperature, SinT, from 550 to 1100 K evaluated from the whole dataset, i.e., from the initial data to the 4th ALMLBO cycle included.

Fig. 3 shows the evolution of the measured zT *mean* values for the kesterite $Cu_{2.125}Zn_{0.875}SnS_4$ before (Init.), during (1rd-4th cycles), and after ALMLBO (Reproducibly Check). The initial (pre-ALMLBO), 1st, 2nd, 3rd, and 4th cycle data are depicted by gray, blue, purple, red, and green squares, respectively. The three remaining white squares denote a reproducibility study with high-purity chemicals performed *a posteriori* of the ALMLBO pipeline. Two dashed red lines illustrate the samples with the highest reported zT *mean* values in this study, i.e., $zT = 0.36$, 0.40@675 K. A quadratic summation of the statistical and systematic uncertainties on the zT is fixed at ~17.5 % [50]. Before using an ALMLBO, zT *mean* values remain in a common and stable domain [0.25, 0.30] that initially seems to be difficult to overcome. Nevertheless, from the 1st cycle of the ALMLBO pipeline, experimental zT *mean* values at 675 K get diverse, spanning from 0.19 to 0.35, expanding by more than a factor 3 the range of expected zT. Passing by two exploration (1st, 3rd cycles) and two exploitation (2nd, 4th cycles) stages, the ALMLBO pipeline reaches an improved $zT = 0.36$@675 K (pre-reproducibility study) with only 15 more experimental samples, allowing a very efficient improvement of the thermoelectric performance of the kesterite $Cu_{2.125}Zn_{0.875}SnS_4$ by checking only ~0.07 % of all the experimentally possible 36,608 combinations of SPS process parameters $\mathbf{\Theta}$. Notably, the ALMLBO approach identified optimal processing parameters that were not immediately obvious from initial experiments or conventional wisdom. For instance, the combination of high heating rate ($\geq$ 150 K/min) and low sintering temperature ($<$ 773 K) led to improved zT, a counterintuitive result that might have been missed in a traditional parameter sweep. Therefore, while the predictions of the ALMLBO approach are valuable, its primary strength lies in guiding the experimental process. The 60 % improvement in zT achieved through this method demonstrates its efficacy in navigating complex parameter spaces, even with limited initial data. **Fig. S7** illustrates the evolution of S, ρ, κ_{tot} as well. Additionally, **Fig. S8** shows the evolution of SPS process parameters $\mathbf{\Theta}$ along the ALMLBO pipeline development, where a regimen of stability has been established for all of $\mathbf{\Theta}$ parameters. Indeed, once sample 26 is evaluated with the highest observed zT *mean*, a plateau for the heating rate, sintering temperature, cooling rate, and uniaxial pressure is clearly reached. The remaining erratic behavior for the step time may reveal its poor relationship to other $\mathbf{\Theta}$ parameters conjugated effects at affecting zT. The reach of this plateau for most of $\mathbf{\Theta}$ parameters during the ALMLBO pipeline is an indicator that a local, potentially global, optima is reached and the ALMLBO process can be stopped.

3.2. *Sintering process archetype*

As the result of the ALMLBO, we can define three archetypes of process parameters (Fig. 4) which induce a significant improvement, or degradation, of the kesterite zT in the temperature range compared to a reference sample (cf. Experimental procedures part). During the ALMLBO step, the heating and cooling rate, temperature, step time, and uniaxial load have been investigated leading to the following conclusions: *i*) in the parameters ranges currently studied, the sintering process

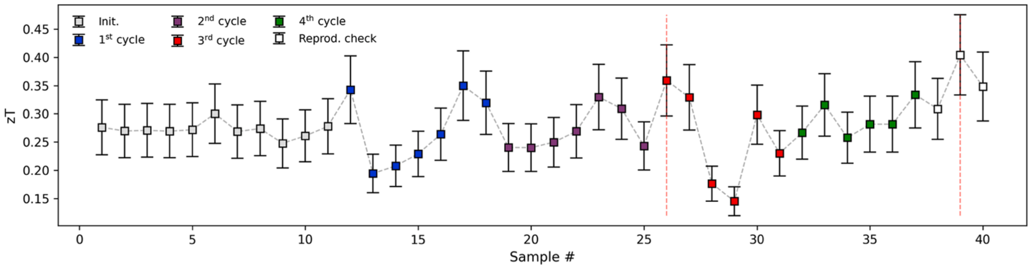


Fig. 3. Evolution of the measured zT *mean* values at 675 K of the kesterite $Cu_{2.125}Zn_{0.875}SnS_4$ along the ALMLBO pipeline development. The initial (pre-ALMLBO), 1st, 2nd, 3rd, and 4th cycle data are depicted by gray, blue, purple, red, and green squares, respectively. The three white squares denote a reproducibility study performed with high-purity chemicals a posteriori of the ALMLBO pipeline. Two dashed red lines illustrate the samples with the highest reported zT values in this study, i.e., $zT = 0.36$, 0.40. Error bars denote the joint statistical and systematic uncertainties on the zT at ~17.5 % [50].

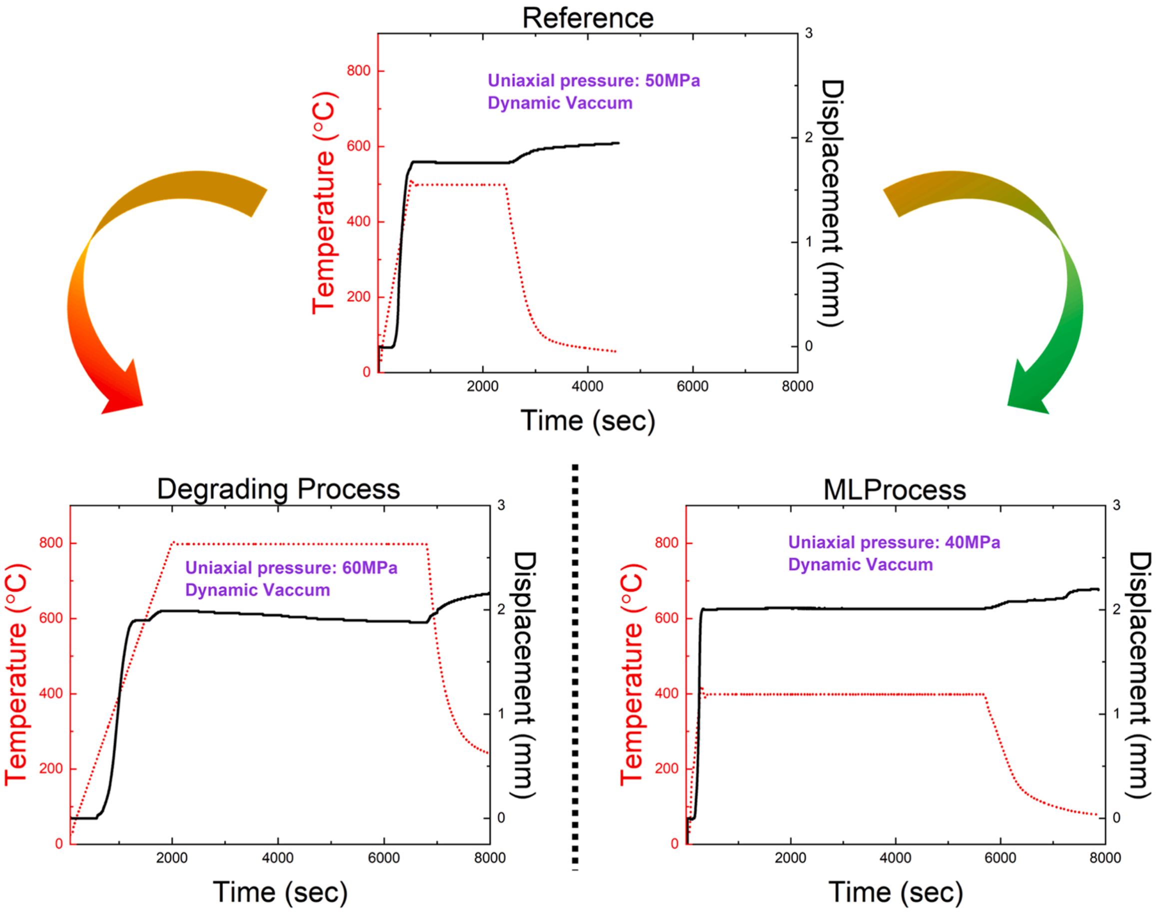


Fig. 4. Representative archetype of SPS process and displacement parameters for Reference, MLProcess and Degrading Process.

is mostly non-reactive as attested by the characteristic displacements curves showing homogenous sintering (Fig. 4), *ii*) high heating rate (≥ 150 K/min), low sintering temperature (< 773 K), low uniaxial load (< 50 MPa) combined with long step time (> 40 min) are favorable to obtain stable and improve zT kesterite (ex. MLP), *iii*) high sintering temperature (> 873 K) and high uniaxial load (≥ 60 MPa) degraded the performance and/or the stability of the kesterite. For this latter one, the low heating rate (< 100 K/min) and long step time (> 30 min) are susceptible to exacerbate the degradation as well as even induce a metastable character of the final kesterite ceramic (ex. DP). As a sign of this degradation, we can observe on the displacement curve of the DP sample that from $T = 873$ K a second sintering shoulder is observable during the heating as well as an anomalous dilation of the sample during the step time at high temperature is visible by comparison with the Ref or MLP samples wherein no anomalous behaviors are observed.

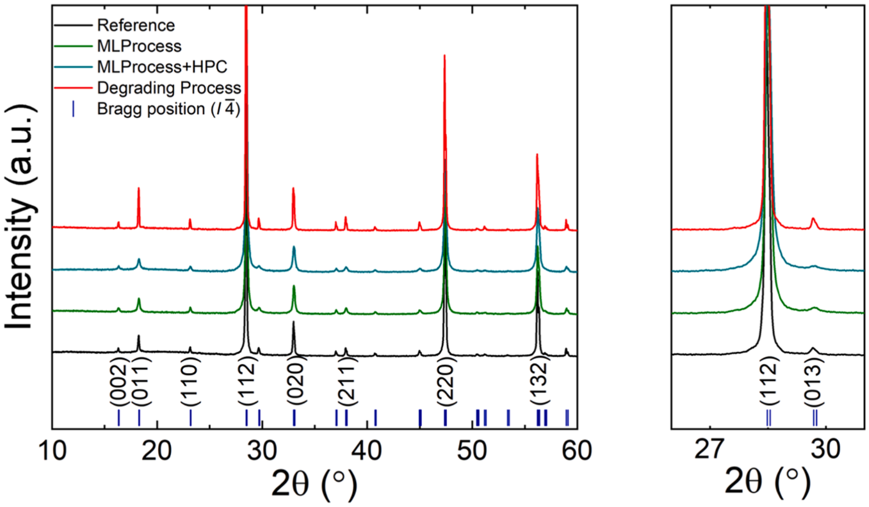


Fig. 5. X-ray diffraction patterns of the CZTS SPSed samples.

3.3. Structural and microstructural analysis

All the different processes led systematically to the formation of the single-ordered phase kesterite (Space group $I\bar{4}$ (n° 82), $a \approx 5.42$ Å, $c \approx 10.83$ Å) as evidenced by the room temperature X-ray powder diffraction (XRPD) pattern displayed in Fig. 5. As the crystal structure of the ordered ($I\bar{4}$) and disordered ($I\bar{4}2m$) kesterite are too similar (**Fig. S9**), the XRPD cannot distinguish the two structures and no conclusion can be withdrawn regarding the possible presence of residual disordered phase induced by the process, however, the ordered structure is the stable form at room temperature of the kesterite [42]. Therefore, the ordered structure has been considered in the Rietveld refinement and we confirmed the phase formation from the low-reliability factor (Table 2 **and Table S2**).

The Rietveld refinement revealed a variation in the R_{Bragg} (intensity-based R-Factors) and R_F (unweighted single-crystal R-factor) [62] factor in the DP sample attesting to some slight structural divergence of this sample compared to the other. It is expected that the DP sample might accumulate structural defects such as a variation of the cationic occupancy or the formation of vacancy due to the long process at high temperature which will affect the average crystal arrangement and therefore will lead to lower reliability factors. Besides, the process does not significantly affect the structural lattice parameters which support a non-variation of the chemical composition. However, thoughtful analysis of the pattern can reveal that the peak broadening is larger in the ML samples (Fig. 5**-right**). The peak full width at half maximum (FWHM) has been estimated on the three major reflections (112), (220), and (132) and evidence the significant increase of the peak broadening (Table 2). In X-ray diffraction, the peak broadening is related to the residual strain (amorphization) within the structure or submicronic crystallite sizes according to the Scherrer equation [63]. In our current study, the conventional and non-reactive process (**cf. sintering process archetype**) make unlikely the strain formation within our material structure and supported a nano-scale size of the crystallite which has been confirmed by the SEM imaging (Fig. 6).

It is well known that mechanical alloying (MA) produces fine powder with a sub-micron particle size which has been confirmed in this present study (**Fig. S10**). The XRPD confirmed that the powder is crystallized, and the main diffraction peak agrees with the kesterite structure (**Fig. S10a)**). The raw powder is composed of an agglomeration of nanoparticles (**Fig. S10.b)**) with an estimated size of the elementary particle clearly below 100 nm with an angular high sphericity shape (**Fig. S10.c)**). The grain sizes and shapes are strongly affected by the sintering procedure as evidenced by the micrograph of the fractured surface of the three representative samples (Fig. 6). The MLP sample is constituted by significantly reduced grain sizes (<< 1 μm) in comparison with the reference sample (2 – 8 μm) agreed with the previous XRPD observation.

According to the classical grain growth kinetics, the average grain size of nanocrystalline materials can be described as [64–67]:

$$D - D_0 = Kt^{1/n} \tag{2}$$

Where D is the mean grain size, D_0 is the initial grain size, K is the temperature-dependent constant, t is the annealing time and n is the grain growth coefficient representing the grain growth behavior. Despite the high kinetic (≥ 150 K/min) and long sintering time (≈ 45 min), the MLP samples present a quasi-non-grain growth in front of the reference which suggests that the grain growth is mainly temperature dependent in the kesterite with a temperature of activation above 673 K (MLP sample) but below 773 K (Reference sample) and its confirmed systematically (**Fig. S11a)**). In another way, the DP sample ($T_{sinter} = 1073$ K) displayed in Fig. 6**c)** present also significantly larger grain size compared to the initial raw nanopowder, however, the average grain size seems close to the reference sample. This support that the grain growth coefficient should be large which implies a extremely fast grain growth kinetic for reaching maximum values around $D_{max} \approx 10$ - 15 μm (**Fig. S11b)**) when the sintering temperature is above the critical activation temperature. This crucial finding enables the possibility to nanostructured kesterite material by knowing the preponderant parameters of its grain growth mechanism. Aside from grain sizes, the microstructure of the MLP and Ref are comparable and present angular spherical grain shapes similar to the initial raw powder sustaining an isotropic feature. The DP sample is characterized by an inhomogeneous microstructure with some areas composed by well define isotropic grain shapes, comparable to the reference (Fig. 6c.3)), and some areas composed of fractured microstructure overall severely compromised (Fig. 6c.1)2)). This attests to the degrading aspect of the process combining high temperature, uniaxial pressure, and step time which produce an unstable final ceramic. This affected drastically the transport properties (cf. thermoelectric properties part) but also the mechanical strength of the bulk which is particularly brittle. Such phenomenon has been observed in other multi-cationic Cu-based sulfides wherein slight structural variation, such as exsolution or sulfur over-evaporation, to cite a few, can conduct to a systematic failure of the microstructure coherence and stability [23,24,33,34,68–71]. Indeed, the Hall-Petch relation described the yield stress ε scales with the inverse square root of grain size δ in polycrystalline materials according to the following relation:

$$\varepsilon = \varepsilon_0 + k \Big/ \sqrt{\delta} \tag{3}$$

where ε_0 and k are chemistry- and microstructure-dependent constants. It highlights that a smaller grain size strengthens the yield stress, as well as a homogeneous grain distribution, limits the intergranular stress [72]. In consequence, the large grain with inhomogeneous sizes distribution of the DP sample is auspicious to microstructure failure under mechanical stress as clearly observed in Fig. 6**c)**.

Interestingly, despite the significant impact of the process on the microstructure, the final composition of the kesterite is barely affected and remain close to the theoretical nominal composition for all the different process according to the SEM-EDS analysis (Table 3). Moreover, the distribution of the elements is homogeneous among the sample even in the DP sample (**Fig. S12**).

3.4. Thermoelectric properties

The transport properties of the representative samples are displayed in Fig. 7. The $|S|$ are positive confirming the p-type character with comparable values independent of the process applied. The monophasic and identical crystal structure (Fig. 5) with no significant chemical composition variations (Table 3) implies similar carrier concentration

Table 2
Rietveld refinement parameter of the X-ray powder diffraction (XRPD) patterns of the CZTS SPSed samples.

$Cu_{2.125}Zn_{0.875}SnS_4$; $\lambda\kappa_{\alpha}Cu = 1.5406$ Å; Space group $I\bar{4}$ (n° 82)							FWHM (°)		
	a (Å)	**c (Å)**	**V(Å^3)**	**χ2**	**R_{Bragg}(%)**	**R_F(%)**	**(112)**	**(220)**	**(132)**
Ref	5.426(1)	10.840(1)	319.11(5)	1.85	3.83	3.81	0.1113	0.1077	0.1084
MLP	5.425(1)	10.835(1)	318.95(5)	2.04	3.47	2.44	0.1176	0.1374	0.1512
MLP+HPC	5.426(1)	10.833(1)	318.97(5)	2.79	5.52	2.94	0.1413	0.1662	0.1849
DP	5.427(1)	10.836(1)	319.19(5)	2.54	6.6	6.28	0.0891	0.0959	0.1011

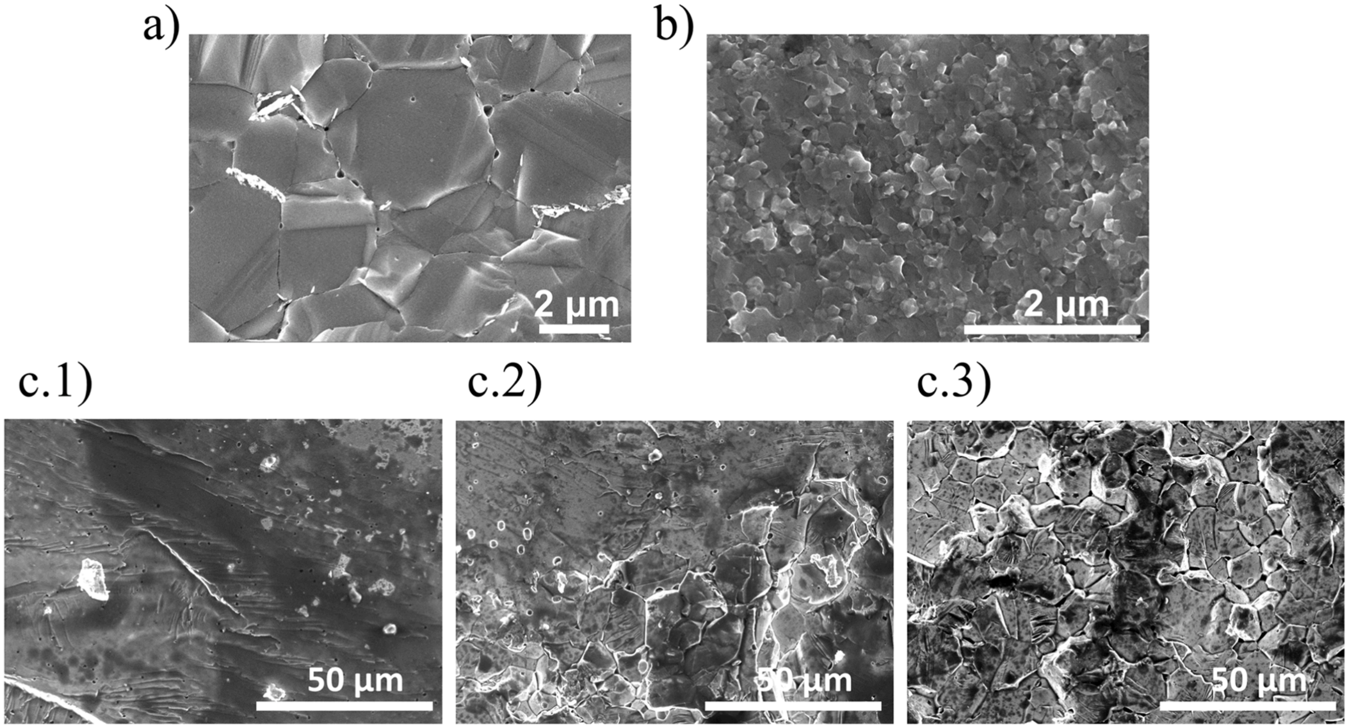


Fig. 6. Scanning electron microscopy (SEM) images of a) reference and b) MLProcess c) Degrading process (where 1), 2) and 3) are different areas) of the CZTS-SPSed samples.

Table 3
Composition analysis by SEM-EDS of the CZTS SPSed samples.

EDS analysis- Elemental atomic percent

	Cu		Zn		Sn		S	
	(at.%)	stdev	(at.%)	stdev	(at.%)	stdev	(at.%)	stdev
Nominal	26.56	–	10.94	–	12.50	–	50.00	–
Reference	27.81	0.84	11.85	0.59	12.61	0.53	47.73	0.97
MLP/MLP+HPC	28.32	0.56	11.42	0.33	12.80	0.31	47.46	0.87
DP	27.92	0.94	11.50	0.52	12.44	0.56	48.14	1.09

and effective mass due to the unchanged band-structure, therefore is expected that the Seebeck coefficient will be similar for the same nominal composition $Cu_{2.125}Zn_{0.825}SnS_4$ despite the different process. This is confirmed through the room temperature Hall effect measurement giving a carrier concentration of $h = 3.94(\pm1.50)\text{x}10^{20}$ cm^{-3} and $3.08(\pm0.90)\text{x}10^{20}$ cm^{-3}, respectively, for the Ref and MLP samples which correspond to a carrier doped CZTS. Besides, the samples have a similar temperature dependence of their Seebeck coefficient, with the typical increasing trend with the temperature of a semiconductor and the characteristic slope change at the (order-disorder) phase transition temperature ($T \approx 500$ K) also observed in the literature [35,36,39,40,43]. The phase transition is also observable on the resistivity curve which exhibits a drop in value at the transition and shows once more that, independently of the process, all samples present a similar temperature dependence. The resistivity increases slightly with temperature in the fully ordered ($T < 425$ K) or disorder ($T > 475$ K) region characteristic of a highly doped semiconductor behavior due to the Cu_{Zn} point defects induced by the native off-stoichiometry $Cu_{2+x}Zn_{1\text{-}x}SnS_4$ selected. The intrinsic substitution of Zn by Cu implies the formation of Cu^{2+} in order to maintain the formal charge balance and, therefore, generate one hole and increase the carrier concentration. As evidence, the stoichiometric Cu_2ZnSnS_4 is reported with a low carrier concentration in the order of $h = 1 \times 10^{17}$ cm^{-3}@300 K by comparison with the current off-stoichiometric sample with values 3 orders of magnitude larger $h \approx 3\text{–}4 \times 10^{20}$ cm^{-3}@300 K typical of a highly doped semiconductor [36]. Similarly to the Seebeck coefficient, the resistivity values are similar between the reference and the MLP samples despite the reduced grain size in the MLP samples indicating that the carrier mobility is not significantly reduced by the carrier scattering at the grain interface. The TE sulfide-based materials are often reported with low carrier mobilities, as in the present case estimated to be $\mu_h = 4.85$ cm^2/V.s and $\mu_h = 4.17$ cm^2/V.s, respectively, for the reference and MLP samples. The holes mobility in this system can be defined simply by:

$$\mu_h = \frac{e\tau_c}{m_h^* V_F} \tag{4}$$

Where μ_h is the hole mobility, e is the elementary charge, τ_c is the mean free path, m_h^* the carrier effective mass, and V_F the Fermi velocity. Assuming that the hole effective mass is comparable considering the similar carrier concentration and Seebeck coefficient, it implies that the mean free path of the hole in this system is equivalent and likely far smaller than the grain size ($\tau_c << D$) considering that it's not affected by the nanostructuring. However, the grain boundaries effect is significant in reducing the thermal transport (Fig. 7d). This striking finding attests that nanostructuring is the key aspect to reach, with the process in our case, to efficiently uncoupling the electrical transport and the thermal transport in kesterite. We can notice that the MLP+HPC sample presents a discrepancy, especially in the room temperature range (Fig. 7**b)**), with

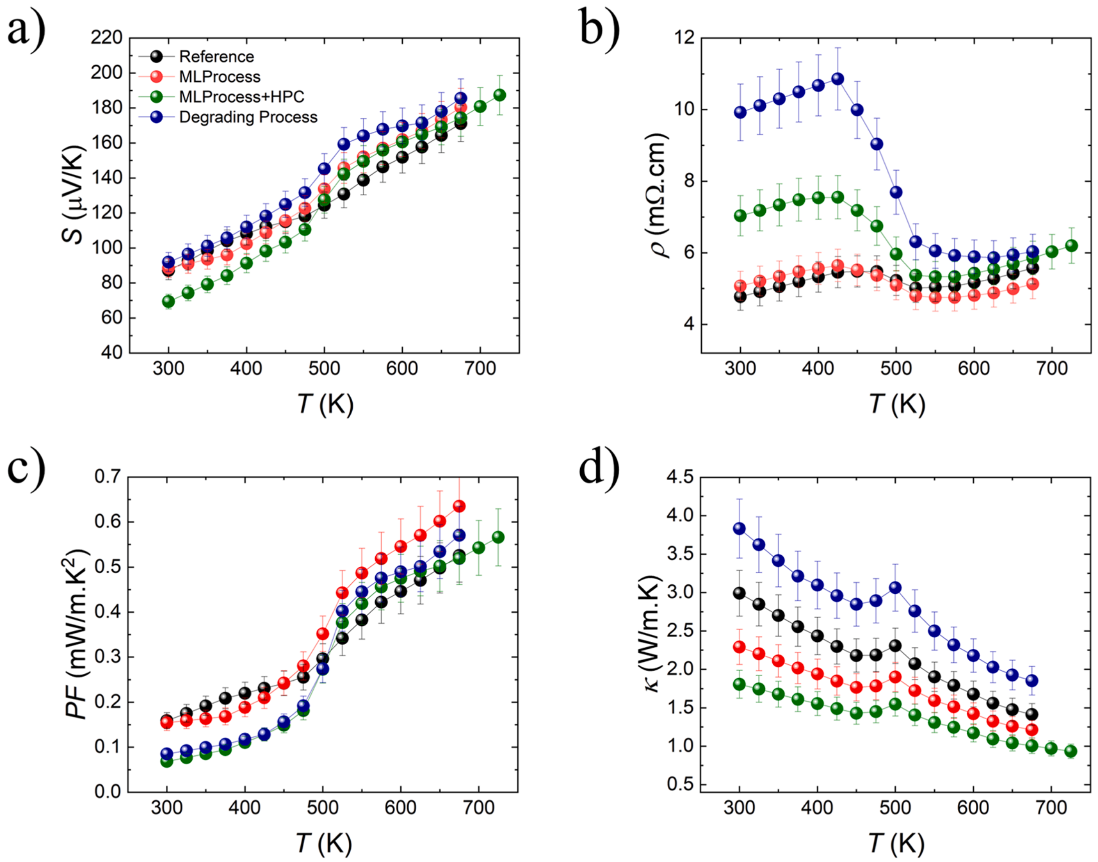


Fig. 7. Dependance in T of a) Seebeck coefficient S, b) electrical resistivity ρ, c) power factor *PF* and d) thermal conductivity κ of the CZTS SPSed samples.

larger ρ related to slightly lower relative density (94.7 %) compared to the reference or MLP sample (97.6 % and 98.5 % respectively). However, it is important to notice that the MLP+HPC presented excellent thermal stability even superior to the reference and MLP samples as evidenced by the reproducibility of the measurement (**Fig. S13**). Finally, the representative DP sample exhibits a larger ρ around room temperature proving that this process, leading to a degraded microstructure (Fig. 6**c**), can extensively affect the electrical transport properties mainly in the ordered region of the CZTS. As a result of the electrical transport properties, the power factor (*PF*) has been determined and displayed in Fig. 7c. All the processes employed led to a similar *PF* trend with a small discrepancy in values. In the disordered region, the *PF* benefit simultaneously from the S bump up and the ρ drop enabling large values lying in the range of $PF = 0.40$ - 0.63 mW/m.K^2. The MLP process presents the best *PF* with a maximum $PF_{max} = 0.63$ mW/mK^2@675 K. It is worth noting that, to the best of our knowledge, this value constitutes the current record on polycrystalline samples of kesterite-based material. In the ordered region, the *PF* is strongly dependent on the process which directly influences the microstructure and therefore the ρ as evidenced by the MLP+HPC and DP samples (Fig. 7b and c) wherein the values are reduced by around 50 %.

The main effect of the process is visible in the thermal conductivity κ displayed in Fig. 7d. The dominant contribution of the κ is the lattice contribution (**Fig. S14.a)**). We can observe a characteristic bump at 500 K correlated to the heat capacity *Cp* due to the phase transition (**Fig. S14.b)**). We can observe that the κ is greatly reduced, in all temperature range, in the MLP samples compared to the reference and reversely increased in the DP samples. To better understand the difference of κ among our samples, we performed phonon transport calculations based on density functional theory (DFT). Calculated phonon dispersion and density of states (DOS) are shown in **Fig. S15**. The dispersion curve consists of three acoustic and 21 optical branches. We included non-analytical term correction to dynamical matrix by calculating dielectric constant and Born effective charge. As a result, longitudinal optical-transverse optical (LO-TO) splitting near the zone center induced by long-range Coulomb interaction can be observed. A frequency gap can be observed in the dispersion and DOS mainly because of the large mass difference between cations (Cu, Zn, Sn) and anion (S). Fig. 8**a)** shows the calculated lattice thermal conductivity of single crystal CZTS for crystallographic *a*-axis (*xx*) and *c*-axis (*zz*) directions together with the experimental ones of our polycrystalline samples. The calculated value is significantly higher than experimental ones, which clearly shows that our polycrystalline samples contain phonon scatterers impeding thermal transport. Fig. 8b) shows the cumulative thermal conductivity κ_{cum} at 300 K as a function of cutoff phonon mean free path (MFP) l. The κ_{cum} is defined as:

$$\kappa_{\text{cum}}(l) = \frac{1}{3}\int_{0}^{l} C v_g L dL \tag{5}$$

where C, v_g, L are phonon specific heat, group velocity, and MFP, respectively. The κ_{cum} gives us information on which scale of phonon is dominant for thermal transport. As seen in colored regions in Fig. 8b**)**, three different regions in terms of the slope of curve can be defined: 1–20 nm, 20–200 nm, and 200–1000 nm. It clearly suggests that the phonons with MFP of 200–1000 nm have the most significant contribution to thermal transport in CZTS. Thus, the lattice thermal conductivity of CZTS is sensitive to the fine control of microstructure within this scale. The ALMLBO approach independently converged on processing parameters that resulted in grain sizes within this crucial 200–1000 nm range, demonstrating its ability to capture and exploit fundamental physical relationships without explicit programming of these constraints. The MLP samples are characterized by a fine microstructure

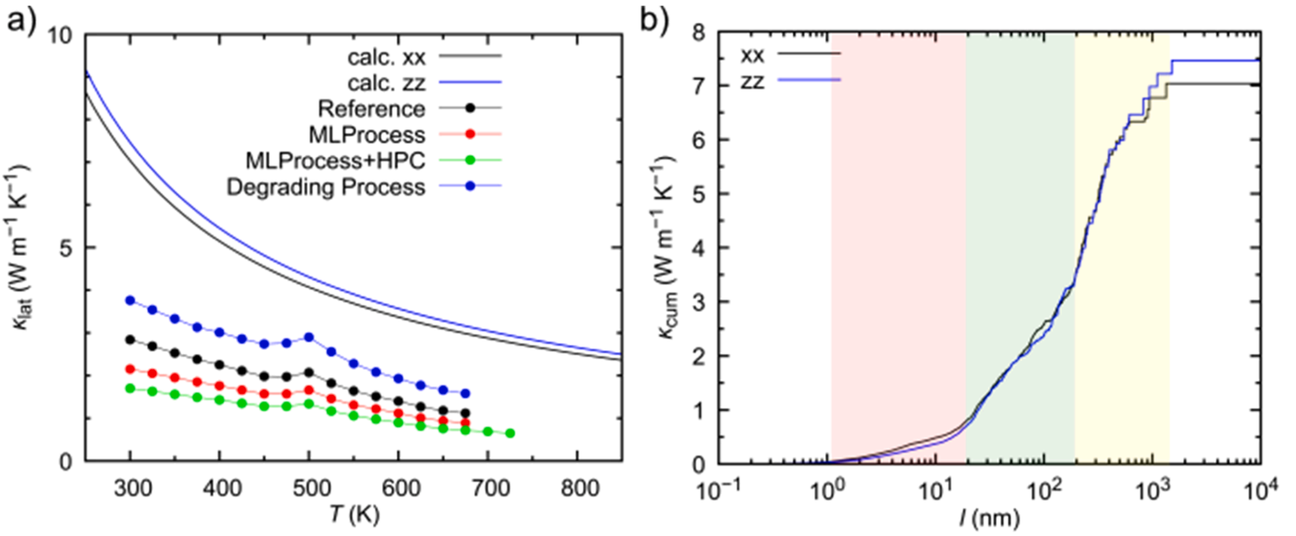


Fig. 8. a) Calculated (solid lines) and experimental (plots) temperature dependence of lattice thermal conductivity of CZTS and b) the cumulative thermal conductivity κ_{cum} at 300 K as a function of the phonon mean free path, l. The colored regions represent that the slope of κ_{cum} curve is significantly different between these regions.

with grain size $D << 1$ µm certainly in the third region of 200–1000 nm which can decrease the κ theoretically up 50 % (Fig. 8b)). Experimentally, we evidence that the κ is reduced by 25 % and 40 % respectively for MLP and MLP+HPC samples at room temperature matching with the calculation. We also show that a bad control of the process can lead to an increase of the κ by 28 % (DP sample). Note that phonons with MFP smaller than 1 nm have negligible contribution to thermal transport in CZTS, which suggests the effect of atomic scale Cu-Zn disorder can be much smaller than the microstructure effect. This finding revealed through the combination of ML-guided experimentation and theoretical calculations, challenges previous assumptions about the relative importance of atomic disorder and microstructure in determining thermal properties of CZTS. We roughly estimated the disorder effect considering only mass difference between Cu and Zn by applying Tamura's model [73] and found that it only gives a small reduction of the lattice thermal conductivity (**Figure S16**).

As illustrated in Fig. 9, the MLP enables a significant improvement by 40–60 % of the zT in the disordered region. It corresponds to a net improvement of the value from $zT \approx 0.27$ *to* $zT \approx 0.36$ at 675 K from the reference to the MLP/MLP+HPC (Fig. 9a)). Also, we performed a two-way mixed ANOVA (analysis of variance) to support the statistical significance of the improvement of the MLP/MLP+HPC on the reference over the whole temperature range [74]. A two-way mixed ANOVA statistical test is here appropriate for our study as it follows the basic assumptions that the different measurements are i.i.d., i.e., independent and identically distributed, are normally distributed, and possess homogeneity of variances, i.e., the fixed variability of 17.5 %, intrinsically implying that their sphericity, i.e., the variances of the differences between all pairs of temperature points are equal. Here, the two-way mixed ANOVA considers the repeated measurements at different temperatures and the comparison between the reference and the MLP/MLP+HPC. The results of the two-way mixed ANOVA show a statistically significant effect of the MLP/MLP+HPC with a p-value $p_0 \sim 9.2 \times 10^{-4}$ and $p_1 \sim 7.1 \times 10^{-3}$ for MLP and MLP+HPC, respectively, indicating that the MLP/MLP+HPC significantly improved the zT values as compared to the reference, averaged across all temperature points. The effect of temperature was also checked for significance for evaluating the consistency of the statistical method on the expected effect of temperature on zT, with a p-value of $p_2 \sim 3.6 \times 10^{-8}$ and $p_3 \sim 3.6 \times 10^{-7}$ for MLP and MLP+HPC, respectively, suggesting significant differences in zT values across different temperatures. The extremely low p-values p_0/p_1 on the MLP/MLP+HPC, much smaller than the commonly used significance level of 0.05, is a strong evidence against the null hypothesis, suggesting that the observed differences between the MLP/MLP+HPC and the reference are unlikely to have occurred by chance alone. Therefore, the improvements in zT obtained from the

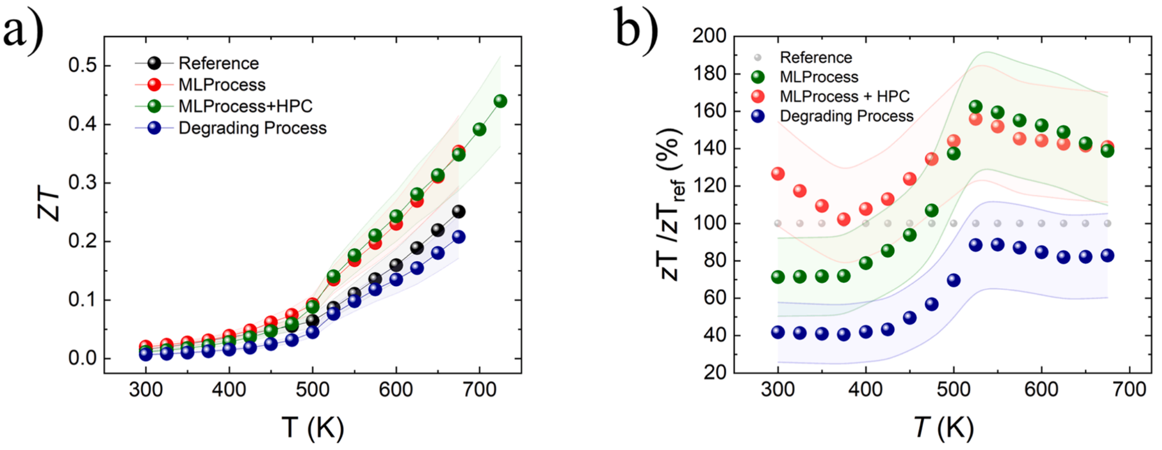


Fig. 9. Dependence in temperature of the a) figure of merit zT of the CZTS SPSed samples. b) improvement ratio normalized on the reference CZTS.

MLP/MLP+HPC are indeed statistically significant across the studied temperature range.

The superior thermal stability of the MLP+HPC sample gives a zT_{max} value of 0.44 at 725 K which corresponds to a record value for the pure kesterite following this process type. In contrast, it highlighted that some process could act negatively on the final figure of merit and led to a degradation of the zT with values reduces by 10–60 % depending to the temperature range (sample DP) evidence once more the importance of the process's control on sulfide-based TE materials. It is worth noticing that the report of Q. Jiang *et al.* shows a peak value at 0.64 at 673 K for a similar composition and process, however, that peak value came mainly from an anomalously low heat capacity which contradicts other literature data and our current measurements (**Fig. S14.b**) [39]. For comparison, using their reported values of *Cp* will make the MLP+HPC sample zT_{max} of 0.81 at 725 K. A study is undergoing to settle the accurate value of *Cp* for CZTS materials in literature.

Besides, it is interesting to observe that the ALMLBO approach applied converge to process's parameter suitable for promoting the nanostructuration which is a common strategy applied on TE materials in order to improve their figure of merit but not systematically effective in every system [75].

4. Conclusions

In the present investigation, we showed the attractiveness of a hybrid study combining an active learning pipeline assisted by machine learning and Bayesian optimization with the "classical" material development strategy in thermoelectricity. It allows optimization in a short time the process, which can strongly influence the thermoelectric performance of metal sulfide-based ceramic and provides unique strategy specific to each system where it applies. As a proof of concept, we investigated in detail the spark plasma sintering densification process on the kesterite $Cu_{2.125}Zn_{0.875}SnS_4$ on its structural, microstructural, and TE properties. We confirmed that the optimization and control of the densification process is critical to unravelling the highest potential on metal sulfide TE ceramics and evidenced that using the appropriate sintering process can allow for the same composition to reach non-negligible enhancement of its zT up to 60 % achieved in a shortened number of experiments. Interestingly, we observed that the ALMLBO converged to a set of parameters suitable for a nanostructuration of kesterite which reduces the thermal conductivity of the material without reducing its overall power factor and led to a record maximum $zT = 0.44$ at 725 K. The ALMLBO approach demonstrated high sensitivity and effectiveness even within the narrow zT range of 0.2–0.4, typical for kesterite systems. This ability to discern and predict small but significant improvements highlights its potential for fine-tuning material properties in constrained optimization scenarios common in materials science. Rather than relying solely on intuition or exhaustive parameter sweeps, ALMLBO provides a data-driven framework for iterative experimentation, representing a paradigm shift in materials optimization. Its true strength lies in guiding experimental design within complex, multi-dimensional parameter spaces, even when working with limited initial data and narrow target value ranges. This approach is particularly valuable in materials science, where experiments are often time-consuming and resource-intensive. By suggesting optimal parameters for each experimental cycle, ALMLBO enables a more thorough and efficient exploration of a material's potential than traditional methods alone. This synergy between machine learning and traditional materials science accelerates the discovery and optimization of high-performance thermoelectric materials, extending the value of ALMLBO beyond mere predictive accuracy. Besides, thanks to the support of theoretical calculations, we evidence and correlate the influence of the microstructure on kesterite in order to uncouple the electronic and phonon transport. This alignment between ML-guided optimization and theoretical phonon calculations demonstrates how data-driven approaches can reinforce and extend our understanding of fundamental physical processes in materials. Reversely, we also identify the parameters which act negatively on performance notably the use of long and high-temperature processes which promote instability and microstructure failure as well as a lower figure of merit. We show the possibility of rationalizing all the results obtained from theoretical, AI, and chemistry points of view. This constitutes novel evidence that combining proficiently the "classical" approach of experiment/theory with data science can allow pushing forward the thermoelectric research on other kesterites and potentially on other thermoelectric materials to a new stage.

CRediT authorship contribution statement

Cédric Bourgès: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Guillaume Lambard:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation. **Naoki Sato:** Writing – review & editing, Validation, Software, Methodology, Formal analysis, Data curation. **Makoto Tachibana:** Writing – review & editing, Visualization, Validation, Methodology, Formal analysis. **Satoshi Ishii:** Writing – review & editing, Validation, Supervision, Resources, Project administration. **Takao Mori:** Writing – review & editing, Visualization, Validation, Supervision, Resources, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by International Center for Young Scientist (ICYS) Fellowship (NIMS, Japan), JST Mirai Program JPMJMI19A1, and JSPS KAKENHI 22K14505. The authors would like to acknowledge NIMS Nanofabrication Facility (NIMS, Japan) for SEM analysis and Material Analysis Station (NIMS, Japan) for the XRD analysis. The computation in this work was performed using the Numerical Materials Simulator at NIMS. (A part of) This work was supported by "Advanced Research Infrastructure for Materials and Nanotechnology in Japan (ARIM)" of the Ministry of Education, Culture, Sports, Science and Technology (MEXT). Proposal Number JPMXP1223NM5166.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.actamat.2024.120342.

References

[1] S. Huang, B. Wang, X. Li, P. Zheng, D. Mourtzis, L. Wang, Industry 5.0 and society 5.0—comparison, complementation and co-evolution, J. Manuf. Syst. 64 (2022) 424–428, https://doi.org/10.1016/j.jmsy.2022.07.010.

[2] H. Park, D. Lee, G. Park, S. Park, S. Khan, J. Kim, W. Kim, Energy harvesting using thermoelectricity for IoT (Internet of Things) and E-skin sensors, J. Phys. Energy 1 (2019) 042001, https://doi.org/10.1088/2515-7655/ab2f1e.

[3] T. Hendricks, T. Caillat, T. Mori, Keynote review of latest advances in thermoelectric generation materials, devices, and technologies 2022, Energies 15 (2022) 7307, https://doi.org/10.3390/en15197307.

[4] I.T. Witting, T.C. Chasapis, F. Ricci, M. Peters, N.A. Heinz, G. Hautier, G.J. Snyder, The thermoelectric properties of bismuth telluride, Adv. Electron. Mater. 5 (2019) 1800904, https://doi.org/10.1002/aelm.201800904.

[5] Z.H. Zheng, X.L. Shi, D.W. Ao, W.D. Liu, M. Li, L.Z. Kou, Y.X. Chen, F. Li, M. Wei, G. X. Liang, P. Fan, G.Q. Lu, Z.G. Chen, Harvesting waste heat with flexible Bi_2Te_3 thermoelectric thin film, Nat. Sustain. 6 (2022) 180–191, https://doi.org/10.1038/s41893-022-01003-6.

[6] H. Jang, J.H. Park, H.S. Lee, B. Ryu, S.D. Park, H.A. Ju, S.H. Yang, Y.M. Kim, W. H. Nam, H. Wang, J. Male, G.J. Snyder, M. Kim, Y.S. Jung, M.W. Oh, Regulating Te vacancies through dopant balancing via excess ag enables rebounding power factor

and high thermoelectric performance in *p*-Type PbTe, Adv. Sci. 8 (2021) 2100895, https://doi.org/10.1002/advs.202100895.

[7] B. Jia, Y. Huang, Y. Wang, Y. Zhou, X. Zhao, S. Ning, X. Xu, P. Lin, Z. Chen, B. Jiang, J. He, Realizing high thermoelectric performance in non-nanostructured *n*-type PbTe, Energy Environ. Sci. 15 (2022) 1920–1929, https://doi.org/10.1039/d1ee03883d.

[8] J. Mao, H. Zhu, Z. Ding, Z. Liu, G.A. Gamage, G. Chen, Z. Ren, High thermoelectric cooling performance of *n*-type Mg_3Bi_2-based materials, Science (1979) 365 (2019) 495–498, https://doi.org/10.1126/science.aax7792.

[9] Z. Liu, W. Gao, H. Oshima, K. Nagase, C.H. Lee, T. Mori, Maximizing the performance of n-type Mg_3Bi_2 based materials for room-temperature power generation and thermoelectric cooling, Nat. Commun. 13 (2022) 1120, https://doi.org/10.1038/s41467-022-28798-4.

[10] O. Caballero-Calero, J.R. Ares, M. Martín-González, Environmentally friendly thermoelectric materials: high performance from inorganic components with low toxicity and abundance in the earth, Adv. Sustain. Syst. 5 (2021) 2100095, https://doi.org/10.1002/adsu.202100095.

[11] B. Zhou, L. Chen, C. Li, N. Qi, Z. Chen, X. Su, X.F. Tang, Significant enhancement in the thermoelectric performance of aluminum-doped ZnO tuned by pore structure, ACS Appl. Mater. Interfaces 12 (2020) 51669–51678, https://doi.org/10.1021/acsami.0c16506.

[12] S. Acharya, B.K. Yu, J. Hwang, J. Kim, W. Kim, High thermoelectric performance of ZnO by coherent phonon scattering and optimized charge transport, Adv. Funct. Mater. 31 (2021) 2105008, https://doi.org/10.1002/adfm.202105008.

[13] J. Yu, K. Chen, F. Azough, D.T. Alvarez-Ruiz, M.J. Reece, R. Freer, Enhancing the thermoelectric performance of calcium cobaltite ceramics by tuning composition and processing, ACS Appl. Mater. Interfaces 12 (2020) 47634–47646, https://doi.org/10.1021/acsami.0c14916.

[14] J. Yu, R. Freer, Calcium cobaltite, a promising oxide for energy harvesting: effective strategies toward enhanced thermoelectric performance, J. Phys. Energy 4 (2022) 022001, https://doi.org/10.1088/2515-7655/ac5172.

[15] Y. He, T. Day, T. Zhang, H. Liu, X. Shi, L. Chen, G.J. Snyder, High thermoelectric performance in non-toxic earth-abundant copper sulfide, Adv. Mater. 26 (2014) 3974–3978, https://doi.org/10.1002/adma.201400515.

[16] A.V. Powell, Recent developments in Earth-abundant copper-sulfide thermoelectric materials, J. Appl. Phys. 126 (2019) 100901, https://doi.org/10.1063/1.5119345.

[17] F.H. Sun, H. Li, J. Tan, L. Zhao, X. Wang, H. Hu, C. Wang, T. Mori, Review of current ZT >1 thermoelectric sulfides, J. Materiomics (2023), https://doi.org/10.1016/j.jmat.2023.05.011.

[18] K. Suekuni, K. Tsuruta, T. Ariga, M. Koyano, Thermoelectric properties of mineral tetrahedrites $Cu_{10}Tr_2Sb_4S_{13}$ with low thermal conductivity, Appl. Phys. Express 5 (2012) 051201, https://doi.org/10.1143/APEX.5.051201.

[19] R. Chetty, A. Bali, R.C. Mallik, Tetrahedrites as thermoelectric materials: an overview, J. Mater. Chem. C 3 (2015) 12364–12378, https://doi.org/10.1039/c5tc02537k.

[20] H. Hu, H.L. Zhuang, Y. Jiang, J. Shi, J.W. Li, B. Cai, Z. Han, J. Pei, B. Su, Z.H. Ge, B. P. Zhang, J.F. Li, Thermoelectric $Cu_{12}Sb_4S_{13}$-based synthetic minerals with a sublimation-derived porous network, Adv. Mater. 33 (2021) 2103633, https://doi.org/10.1002/adma.202103633.

[21] K. Suekuni, F.S. Kim, H. Nishiate, M. Ohta, H.I. Tanaka, T. Takabatake, High-performance thermoelectric minerals: colusites $Cu_{26}V_2M_6S_{32}$ (M = Ge, Sn), Appl. Phys. Lett. 105 (2014) 132107, https://doi.org/10.1063/1.4896998.

[22] C. Bourgès, Y. Bouyrie, A.R. Supka, R. Al Rahal Al Orabi, P. Lemoine, O. Lebedev, M. Ohta, K. Suekuni, V. Nassif, V. Hardy, R. Daou, Y. Miyazaki, M. Fornari, E. Guilmeau, High-performance thermoelectric bulk colusite by process controlled structural disordering, J. Am. Chem. Soc. 140 (2018) 2186–2195, https://doi.org/10.1021/jacs.7b11224.

[23] G. Guélou, P. Lemoine, B. Raveau, E. Guilmeau, Recent developments in high-performance thermoelectric sulphides: an overview of the promising synthetic colusites, J. Mater. Chem. C 9 (2021) 773–795, https://doi.org/10.1039/d0tc05086e.

[24] P. Kamińska, C. Bourgès, R. Chetty, D. Gutiérrez-Del-Río, P. Śpiewak, W. Święszkowski, T. Nishimura, T. Mori, Insight into the preponderant role of the lattice size in Sn-based colusites for promoting a high power factor, J. Mater. Chem. A 10 (2022) 10701–10714, https://doi.org/10.1039/d2ta01210c.

[25] P. Qiu, T. Zhang, Y. Qiu, X. Shi, L. Chen, Sulfide bornite thermoelectric material: a natural mineral with ultralow thermal conductivity, Energy Environ. Sci. 7 (2014) 4000–4006, https://doi.org/10.1039/c4ee02428a.

[26] S.O.J. Long, A.V. Powell, P. Vaqueiro, S. Hull, High thermoelectric performance of bornite through control of the Cu(II) content and vacancy concentration, Chem. Mater. 30 (2018) 456–464, https://doi.org/10.1021/acs.chemmater.7b04436.

[27] J. Li, Q. Tan, J.F. Li, Synthesis and property evaluation of $CuFeS_{2-x}$ as earth-abundant and environmentally-friendly thermoelectric materials, J. Alloys Compd. 551 (2013) 143–149, https://doi.org/10.1016/j.jallcom.2012.09.067.

[28] N. Tsujii, T. Mori, Y. Isoda, Phase stability and thermoelectric properties of $CuFeS_2$-based magnetic semiconductor, J. Electron. Mater. 43 (2014) 2371–2375, https://doi.org/10.1007/s11664-014-3072-y.

[29] N. Sato, P.S. Gan, N. Tsujii, T. Mori, Effect of microstructure on lattice thermal conductivity of thermoelectric chalcopyrite $CuFeS_2$: experimental and computational studies, Appl. Phys. Express 14 (2021) 087002, https://doi.org/10.35848/1882-0786/ac1231.

[30] D. Zhang, B. Zhang, Z. Zhou, K. Peng, H. Wu, H. Wang, G. Wang, G. Han, G. Wang, X. Zhou, X. Lu, Ultralow lattice thermal conductivity of cubic $CuFeS_2$ induced by atomic disorder, Chem. Mater. 33 (2021) 9795–9802, https://doi.org/10.1021/acs.chemmater.1c03785.

[31] K. Hashikuni, K. Suekuni, H. Usui, M. Ohta, K. Kuroki, T. Takabatake, High power factor in thiospinels $Cu_2TrTi_3S_8$ (Tr= Mn, Fe, Co, Ni) arising from TiS_6 octahedron network, Appl. Phys. Lett. 109 (2016) 182110, https://doi.org/10.1063/1.4966955.

[32] C. Bourgès, V. Pavan Kumar, H. Nagai, Y. Miyazaki, B. Raveau, E. Guilmeau, Role of cobalt for titanium substitution on the thermoelectric properties of the thiospinel $CuTi_2S_4$, J. Alloys Compd. 781 (2019) 1169–1174, https://doi.org/10.1016/j.jallcom.2018.12.102.

[33] K. Hashikuni, K. Suekuni, H. Usui, R. Chetty, M. Ohta, K. Kuroki, T. Takabatake, K. Watanabe, M. Ohtaki, Thermoelectric properties and electronic structures of $CuTi_2S_4$ thiospinel and its derivatives: structural design for spinel-related thermoelectric materials, Inorg. Chem. 58 (2019) 1425–1432, https://doi.org/10.1021/acs.inorgchem.8b02955.

[34] C. Bourgès, B. Srinivasan, B. Fontaine, P. Sauerschnig, A. Minard, J.F. Halet, Y. Miyazaki, D. Berthebaud, T. Mori, Tailoring the thermoelectric and structural properties of Cu-Sn based thiospinel compounds $[CuM_{1+x}Sn_{1-x}S_4$ (M = Ti, V, Cr, Co)], J. Mater. Chem. C 8 (2020) 16368–16383, https://doi.org/10.1039/d0tc04393a.

[35] E. Isotta, N.M. Pugno, P. Scardi, Nanostructured kesterite (Cu_2ZnSnS_4) for applications in thermoelectric devices, Powder Diffr. 34 (2019) S42–S47, https://doi.org/10.1017/S0885715619000277.

[36] E. Isotta, B. Mukherjee, C. Fanciulli, N.M. Pugno, P. Scardi, Order-disorder transition in Kesterite Cu_2ZnSnS_4: thermopower enhancement via electronic band structure modification, J. Phys. Chem. C 124 (2020) 7091–7096, https://doi.org/10.1021/acs.jpcc.0c00886.

[37] C. Bourgès, R. Al Rahal Al Orabi, Y. Miyazaki, Off-stoichiometry effect on thermoelectric properties of the new p-type sulfides compounds Cu_2CoGeS_4, J. Alloys Compd. 826 (2020) 154240, https://doi.org/10.1016/j.jallcom.2020.154240.

[38] D. Zhang, J. Yang, Q. Jiang, Z. Zhou, X. Li, J. Xin, A. Basit, Y. Ren, X. He, Multi-cations compound Cu_2CoSnS_4: DFT calculating, band engineering and thermoelectric performance regulation, Nano Energy 36 (2017) 156–165, https://doi.org/10.1016/j.nanoen.2017.04.027.

[39] Q. Jiang, H. Yan, Y. Lin, Y. Shen, J. Yang, M.J. Reece, Colossal thermoelectric enhancement in $Cu_{2+x}Zn_{1-x}SnS_4$ solid solution by local disordering of crystal lattice and multi-scale defect engineering, J. Mater. Chem. A 8 (2020) 10909–10916, https://doi.org/10.1039/d0ta01595d.

[40] A. Nagaoka, K. Yoshino, T. Masuda, T.D. Sparks, M.A. Scarpulla, K. Nishioka, Environmentally friendly thermoelectric sulphide Cu_2ZnSnS_4 single crystals achieving a 1.6 dimensionless figure of merit ZT, J. Mater. Chem. A 9 (2021) 15595–15604, https://doi.org/10.1039/d1ta02978a.

[41] J. Li, K. Sun, X. Yuan, J. Huang, M.A. Green, X. Hao, Emergence of flexible kesterite solar cells: progress and perspectives, Npj Flex 7 (2023), https://doi.org/10.1038/s41528-023-00250-7.

[42] C.J. Bosson, M.T. Birch, D.P. Halliday, K.S. Knight, A.S. Gibbs, P.D. Hatton, Cation disorder and phase transitions in the structurally complex solar cell material Cu_2ZnSnS_4, J. Mater. Chem. A 5 (2017) 16672–16680, https://doi.org/10.1039/c7ta03603e.

[43] E. Isotta, C. Fanciulli, N.M. Pugno, P. Scardi, Effect of the order-disorder transition on the Seebeck coefficient of nanostructured thermoelectric Cu_2ZnSnS_4, Nanomaterials 9 (2019) 762–773, https://doi.org/10.3390/nano9050762.

[44] L.M. Antunes, J.J.Plata Vikram, A.V. Powell, K.T. Butler, R. Grau-Crespo, Machine learning approaches for accelerating the discovery of thermoelectric materials, in: ACS Symposium Series, New York, 2022, pp. 1–32, https://doi.org/10.1021/bk-2022-1416.ch001.

[45] X. Wang, Y. Sheng, J. Ning, J. Xi, L. Xi, D. Qiu, J. Yang, X. Ke, A critical review of machine learning techniques on thermoelectric materials, J. Phys. Chem. Lett. 14 (2023) 1808–1822, https://doi.org/10.1021/acs.jpclett.2c03073.

[46] G. Lambard, T.T. Sasaki, K. Sodeyama, T. Ohkubo, K. Hono, Optimization of direct extrusion process for Nd-Fe-B magnets using active learning assisted by machine learning and Bayesian optimization, Scr. Mater. 209 (2022) 114341, https://doi.org/10.1016/j.scriptamat.2021.114341.

[47] J. Rodríguez-Carvajal, Recent advances in magnetic structure determination by neutron powder diffraction, Phys. B Condens. Matter. 192 (1993) 55–69, https://doi.org/10.1016/0921-4526(93)90108-I.

[48] T. Roisnel, J. Rodríguez-Carvajal, WinPLOTR: a windows tool for powder diffraction pattern analysis, Mater. Sci. Forum. 378-381 (2001) 118–123, https://doi.org/10.4028/www.scientific.net/msf.378-381.118.

[49] P. Thompson, D.E. Cox, J.B. Hastings, Rietveld refinement of debye-scherrer synchrotron X-ray data from Al_2O_3, J. Appl. Crystallogr. 20 (1987) 79–83, https://doi.org/10.1107/S0021889887087090.

[50] E. Alleno, D. Bérardan, C. Byl, C. Candolfi, R. Daou, R. Decourt, E. Guilmeau, S. Hébert, J. Hejtmanek, B. Lenoir, P. Masschelein, V. Ohorodnichuk, M. Pollet, S. Populoh, D. Ravot, O. Rouleau, M. Soulier, Invited Article: a round robin test of the uncertainty on the measurement of the thermoelectric dimensionless figure of merit of $Co_{0.97}Ni_{0.03}Sb_3$, Rev. Sci. Instrum. 86 (2015) 011301, https://doi.org/10.1063/1.4905250.

[51] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G.L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A.P. Seitsonen, A. Smogunov, P. Umari, R.M. Wentzcovitch, QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials, J. Phys. Cond. Matter 21 (2009) 395502, https://doi.org/10.1088/0953-8984/21/39/395502.

[52] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. Buongiorno Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A. Dal Corso, S. De Gironcoli, P. Delugas, R.A. Distasio, A. Ferretti, A. Floris, G. Fratesi, G. Fugallo, R. Gebauer, U. Gerstmann, F. Giustino, T. Gorni, J. Jia, M. Kawamura, H.Y. Ko, A. Kokalj, E. Kücükbenli, M. Lazzeri, M. Marsili, N. Marzari, F. Mauri, N.L. Nguyen, H.V. Nguyen, A. Otero-De-La-Roza, L. Paulatto, S. Poncé, D. Rocca, R. Sabatini, B. Santra, M. Schlipf, A.P. Seitsonen, A. Smogunov, I. Timrov, T. Thonhauser, P. Umari, N. Vast, X. Wu, S. Baroni, Advanced capabilities for materials modelling with Quantum ESPRESSO, J. Phys. Cond. Matter 29 (2017) 465901, https://doi.org/10.1088/1361-648X/aa8f79.

[53] P.E. Blöchl, Projector augmented-wave method, Phys. Rev. B 50 (1994) 17953–17979, https://doi.org/10.1103/PhysRevB.50.17953.

[54] J.P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple, Phys. Rev. Lett. 77 (1996) 3865–3868, https://doi.org/10.1103/PhysRevLett.77.3865.

[55] T. Tadano, Y. Gohda, S. Tsuneyuki, Anharmonic force constants extracted from first-principles molecular dynamics: applications to heat transfer simulations, J. Phys. Cond. Matter 26 (2014) 225402, https://doi.org/10.1088/0953-8984/26/22/225402.

[56] L. Breiman, Random Forests, Mach. Learn. 45 (2001) 5–32, https://doi.org/10.1023/A:1010933404324.

[57] B. Efron, The Jackknife, the Bootstrap and Other Resampling Plans, Soc. Indust. Appl. Mathematics (1982), https://doi.org/10.1137/1.9781611970319.

[58] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, É. Duchesnay, Scikit-learn: machine Learning in Python, J. Mach. Learn. Res. 12 (2011) 2825–2830. http://scikit-learn.sourceforge.net.

[59] T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning Data Mining, Inference, and Prediction, 2nd Edition, Springer, 2009.

[60] G. Louppe, Understanding random forests: from theory to practice, 2014.

[61] B. Gregorutti, B. Michel, P. Saint-Pierre, Correlation and variable importance in random forests, Stat Comput 27 (2017) 659–678, https://doi.org/10.1007/s11222-016-9646-1.

[62] B.H. Toby, R factors in Rietveld analysis: how good is good enough? Powder. Diffr. 21 (2006) 67–70, https://doi.org/10.1154/1.2179804.

[63] A.L. Patterson, The scherrer formula for i-ray particle size determination, 1939.

[64] H. Gleiter, Nanocrystalline materials, Prog. Mater. Sci. 33 (1989) 223–315, https://doi.org/10.1016/0079-6425(89)90001-7.

[65] K. Lu, Nanocrystalline metals crystallized from amorphous solids: nanocrystallization, structure, and properties, Mater. Sci. Eng. R: Rep. 16 (1996) 161–221, https://doi.org/10.1016/0927-796X(95)00187-5.

[66] H.F. J, H. M, Recrystallization and Related Annealing Phenomena, 2nd edition, 2004.

[67] K.N. Zhu, Q. Ruan, A. Godfrey, The kinetics of grain growth in near-micrometre grain size copper produced by spark plasma sintering, IOP Conf. Ser. Mater. Sci. Eng. (2015) 012060, https://doi.org/10.1088/1757-899X/89/1/012060.

[68] C. Bourgès, M. Gilmas, P. Lemoine, N. Mordvinova, O.I. Lebedev, E. Hug, V. M. Nassif, B. Malaman, R. Daou, E. Guilmeau, Structural analysis and thermoelectric properties of mechanically alloyed colusites, J. Mater. Chem. C 4 (2016) 7455–7463, https://doi.org/10.1039/C6TC02301K.

[69] T. Barbier, P. Lemoine, S. Gascoin, O.I. Lebedev, A. Kaltzoglou, P. Vaqueiro, A. V. Powell, R.I. Smith, E. Guilmeau, Structural stability of the synthetic thermoelectric ternary and nickel-substituted tetrahedrite phases, J. Alloys Compd. 634 (2015) 253–262, https://doi.org/10.1016/j.jallcom.2015.02.045.

[70] X. Lu, D.T. Morelli, Y. Xia, F. Zhou, V. Ozolins, H. Chi, X. Zhou, C. Uher, High performance thermoelectricity in earth-abundant compounds based on natural mineral tetrahedrites, Adv. Energy Mater. 3 (2013) 342–348, https://doi.org/10.1002/aenm.201200650.

[71] Y. Bouyrie, C. Candolfi, A. Dauscher, B. Malaman, B. Lenoir, Exsolution process as a route toward extremely low thermal conductivity in $Cu_{12}Sb_{4-x}Te_xS_{13}$ tetrahedrites, Chem. Mater. 27 (2015) 8354–8361, https://doi.org/10.1021/acs.chemmater.5b03785.

[72] Z.C. Cordero, B.E. Knight, C.A. Schuh, Six decades of the Hall–Petch effect – a survey of grain-size strengthening studies on pure metals, Inter. Mater. Rev. 61 (2016) 495–512, https://doi.org/10.1080/09506608.2016.1191808.

[73] S.I. Tamura, Isotope scattering of dispersive phonons in Ge, Phys. Rev. B 27 (1983) 858, https://doi.org/10.1103/PhysRevB.27.858.

[74] S.E. Maxwell, H.D. Delaney, K. Kelley, Designing Experiments and Analyzing Data, 3rd edition, Routledge, New York, 2017, https://doi.org/10.4324/9781315642956.

[75] M.G. Kanatzidis, Nanostructured thermoelectrics: the new paradigm? Chem. Mater. 22 (2010) 648–659, https://doi.org/10.1021/cm902195j.