

Closed-loop Error Correction Learning Accelerates Experimental Discovery of Thermoelectric Materials

Hitarth Choubisa[†], Md Azimul Haque[†], Tong Zhu, Lewei Zeng, Maral Vafaie, Derya Baran^{}, Edward H Sargent^{*}*

Dr. H. Choubisa, Dr. Tong Zhu, L. Zeng, Dr. M. Vafaie, Prof. E. H. Sargent
Department of Electrical and Computer Engineering University of Toronto, Toronto, Ontario, Canada
ted.sargent@utoronto.ca

Dr. M. A. Haque, Prof. D. Baran
King Abdullah University of Science and Technology (KAUST), Physical Science and Engineering Division, KAUST Solar Center (KSC), Thuwal, Saudi Arabia
derya.baran@kaust.edu.sa

[†] H.C. and M.A.H. contributed equally to this paper.

Keywords: Machine learning, thermoelectrics, closed-loop, error-correction learning

Abstract:

The exploration of thermoelectric materials is challenging considering the large materials space, combined with added exponential degrees of freedom coming from doping and the diversity of synthetic pathways. Here we seek to incorporate historical data and update it using experimental feedback by employing error-correction learning (ECL). We thus learn from prior datasets and then adapt the model to differences in synthesis and characterization that are otherwise difficult to parameterize. We then apply this strategy to discovering thermoelectric materials, where we prioritize synthesis at temperatures $< 300^{\circ}\text{C}$. We document a previously-unexplored chemical family of thermoelectric materials, PbSe:SnSb, finding that the best candidate in this chemical family, 2 wt% SnSb doped PbSe, exhibits a power factor more than 2x that of PbSe. The investigations herein reveal that a closed-loop experimentation strategy reduces the required number of experiments to find an optimized material by a factor as high as 3x compared to high-throughput searches powered by state-of-

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1002/adma.202302575](https://doi.org/10.1002/adma.202302575).

This article is protected by copyright. All rights reserved.

art machine learning (ML) models. We also observe that this improvement is dependent on the accuracy of the ML model in a manner that exhibits diminishing returns: once a certain accuracy is reached, factors that are instead associated with experimental pathways begin to dominate trends.

1. Introduction

Thermoelectric materials convert thermal energy to electricity based as described using the Seebeck effect, wherein a thermal gradient leads to the generation of electric current¹. Their capacity to convert waste heat to electricity provides a route to recover the energy lost in mechanical and electrical processes². Today's best-performing thermoelectric materials are particularly effective at high temperatures; and thus further progress is necessary to enable energy harvesting in low-temperature consumer applications such as IoT devices and wearables^{3,4}. The design and discovery of thermoelectric materials is challenging in view of the large chemical space^{5,6}, non-convex composition-property mapping⁷ and nonlinear effects of dopants on material properties^{8,9}.

Large-scale chemical space searches for materials with desired properties have been pursued using high-throughput Density Functional Theory (DFT) based simulations^{10–17} and machine learning (ML) based property prediction searches^{8,18–21}. Many of these studies rely on high-throughput DFT screening to rank material candidates; after which the most promising candidates are then tested in the lab^{22–24}. However, the large computational cost associated with DFT calculations prohibits extensive exploration of the chemical space; and this approach does not account for the gap between simulations and experiments. The former drawback has been addressed in recent years with the help of machine learning models that learn from existing datasets and screen much larger chemical spaces^{25–27}. The inability to bridge the gap between theory and experiments has been addressed in a small number of studies; valuable precedents in literature achieve this using closed-loop strategies^{28–34}.

Prior such studies either required pre-defined functional forms for the data fusion methods to integrate available data and acquired data; or the presence of relevant sampled data to estimate the property distribution. The first requirement limits how information can be transferred between the two data sources; while the latter is computationally expensive²⁸ or may even be impossible to curate from available datasets. Consequently, accurate learning of material-property relationships is limited, and this militates against efficient materials

exploration quantified by accuracy scores on data acquired and ranked according to performance.

We hypothesized that a data-driven strategy for information transfer between the available and acquired data could overcome the first limitation. The problem of uniformly sampling the materials space to acquire an informed prior can then be addressed independently: it becomes equivalent to the development of accurate ML surrogate models. We combine these two strategies, and demonstrate a two-step error correction learning approach: learning from existing datasets and iterative refinements with new experimental results (**Figure 1** for details of the discovery process).

We apply this strategy to the discovery of earth-abundant and low-temperature thermoelectric materials (300K) via optimization of the experimental power factors (PF). We explore the material space through the synthesis of new inorganic compounds, finding that doping and alloying lead to the family of low-temperature thermoelectric materials PbSe:SnSb. The predicted best candidate exhibits a power factor 2x that of PbSe, and is the highest among previously-reported low-temperature-(< 300°C)-synthesized materials. In the process, we find that chemical representation as well as the accuracy of ML models on available datasets play quantitatively significant roles in accelerating closed-loop materials discovery. Exploration using DFT of the origins of higher power factor shows the interaction of Se and Sb in PbSe:SnSb leads to a reduced hole effective mass and increased power factors. The work shows that ECL-based closed-loop approaches account for factors traditionally difficult to parameterize in materials discovery.

2. Results and Discussion

2.1. Error-correction learning with experiments

In a standard ML-based materials screening pipeline, one learns the mapping between material structures/compositions and the outcome variable(s) of interest^{35–38}. The trained model is used to screen large chemical spaces and rank the promising candidates. The top-ranked candidates are then validated experimentally^{8,39}. However, synthesis conditions such as the synthesis method (hydrothermal or sputtering or ball milling), morphologies (powder or single crystal or nanoparticles), post-synthesis treatment (annealing or etching), state of the precursors (metal salts or metal alkoxides) and source of energy (microwave or ultrasound or laser) differ from one experimental laboratory to another. These potentially significant factors

are difficult to parameterize, making it hard to predict synthetic outcomes using ML models trained on historical data, especially from lab to lab. As a result, additional experiments within a new lab are required to explore the chemical space and find the optimal material candidate.

With this in mind, we formulate the challenge as an error-correction problem: the observed outcomes of our in-lab experiments (denoted by y) for two compositions c_1 and c_2 can be expressed as:

$$y = \begin{cases} 0 & \text{if } PF(c_1, T_1) \geq PF(c_2, T_2) \\ 1 & \text{if } PF(c_1, T_1) < PF(c_2, T_2) \end{cases} \#(1)$$

where c_1 and c_2 are the materials we wish to compare, T_1 and T_2 are the temperatures at which we make the comparison, and $PF(\cdot)$ represents the in-lab measured power factor (or generally any other property of interest). y is dependent on the synthesis parameters and methods. It is thus not easy to merge y with existing datasets coherently and train the ML model on the combined data. Instead, we propose to model the observed outcome y using an error-correction strategy,

$$y \sim g_\theta(X_1, X_2, T_1, T_2, m) \#(2)$$

where X_1 and X_2 are vectorial representations for the two materials compositions c_1 and c_2 to be compared; T_1 and T_2 are the temperatures at which to make the comparison; $PF(\cdot)$ represents the power factor (or other property of interest) and m is the *prior model* trained on existing datasets as one would do in a high-throughput ML search pipeline. The exact form of the error correction function is unknown. To avoid parametric biases, we use a dense neural network, g_θ that can learn from data to model the error correction function (**Figure 1a**). The proportion of error correction that is achievable depends on the quantity of data, model m , and methods used for the vectorial representation of chemical compositions.

We hypothesize that error-correction learning (ECL) enables correction of the predictions made by the ML model trained on a historical dataset. The neural network g_θ learns to map the experimental quantities obtained under a certain set of consistent experimental conditions to the new set of experimental conditions. By choosing a neural network (or generic machine learning model with good accuracy), we learn implicitly to account for the factors that have led to changes in experimental observations between historical data and our experiments.

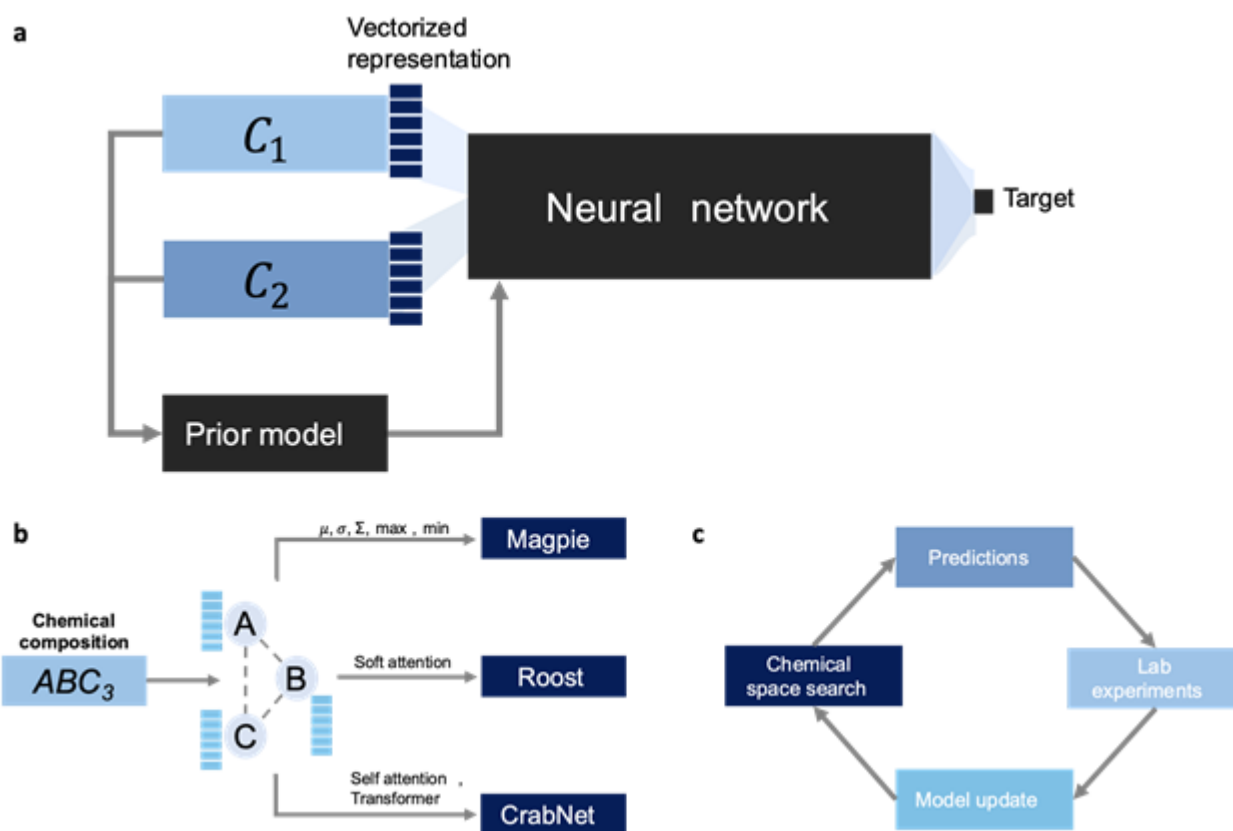


Figure 1. Error correction learning to enable closed-loop materials discovery. (a) The approach taken herein to performing error-correction learning (ECL) on predictions from a pre-trained machine learning model, addressed as the *Prior model*. The error-correction network takes two chemical compositions (C_1 and C_2) as inputs and predicts if C_1 has a power factor larger than C_2 . (b) Three approaches employed and compared for the purpose of vectorial representation of chemical compositions. Magpie⁴⁰ generates vectorial representation using basic statistical operations such as summation, average, standard deviation, min and max operations on elemental properties. Roost³⁸ uses a soft-attention based mechanism and converts stoichiometric graphs to numerical vectorial representations. CrabNet³⁷ generates a numerical representation for chemical compositions using self-attention and transformer architecture. We benchmark all these approaches using their accuracy on literature data and accuracy on collected experimental data. (c) shows our closed-loop approach: we start with the model and make predictions within the chemical space. The top candidates are used for lab experimentation and fed back to our error correction module for improving prediction accuracy.

2.2. Discovery of large power factor thermoelectrics

We apply error-correction learning to the discovery of promising thermoelectric materials. We pursue it by first training a *prior model* on existing data. For training we use a dataset of

thermoelectric materials and their experimental properties⁴¹ compiled in a recent study⁸. The dataset reports thermoelectric metrics: conductivity (σ), figure-of-merit (zT) and power factor (PF), measured experimentally in the lab for 573 doped and alloyed materials, spanning 52 distinct elements. These elements are highlighted in the periodic table of **Figure S1**. The distribution of the material compositions and reported power factors within this dataset are visualized in **Figure S2**. The dataset covers a diverse set of chemical compounds comprising oxides, other chalcogenides, silicides, germanides, metal alloys and antimonides.

We train ML regression models to predict the PFs using material compositions and measurement temperatures as predictors. We use the TPOT library to perform model search, optimal preprocessing step selection and hyperparameter tuning⁴². TPOT searches through various ML models (support vector machines, linear regressions, gradient boosted trees, random forests, decision trees, gaussian processes) and their combinations to find the model with the best performance. For training, the entire dataset is randomly split between 80% training and 20% testing dataset. Hyperparameter optimizations are performed using 5-fold cross-validation seeking to maximize the cross-validation coefficient of determination (r^2) between target and predicted values. Furthermore, since material compositions can be represented using various methods, we test three distinct representations - Magpie⁴⁰, Roost³⁸ and CrabNet³⁷ and choose the one with the best cross-validation score as the prior model for driving closed-loop experiments. **Figure 1b** shows a high-level summary of how these three representations differ. We use transfer learning for training with Roost and CrabNet approaches due to the relatively smaller size of the dataset. The features extracted after the global pooling layer are used to represent different compositions and as input features to train the model (refer to **SI notes 1-2** for details on ML model training and the models considered).

Interestingly, the *prior model* trained using Magpie showed the best cross-validation performance outperforming Roost and CrabNet (refer to **Figure 2a** for a comparison of various representations and ML models we trained and compared). We explore this observation by analyzing the distribution of the generated representations by each of the methods (SI **Figure S3**). Due to its min-max operators, we observe that Magpie generates a less smooth representation that can better capture the doping and alloying effects. On the other hand, Roost and CrabNet have smoother changes in the representation vectors for different materials and that is reflected through lower accuracies for power factor prediction, a property sensitive to small amounts of alloying and doping. Furthermore, to establish the efficacy of TPOT in finding the optimal ML models, we also separately train and optimize

the hyperparameters of the Random Forest regression model using the same training-testing splits and same number of cross-validation splits (**Figure 2a** and **SI notes 1-2**).

After training the *prior model*, we explore the chemical space to discover low-temperature large power factor materials. In exploration, we limit ourselves to specific precursors and elements. We focused on materials synthesized below $<300^{\circ}\text{C}$ and avoided regulated materials from the list {As, Cd, Hg, Te, Tl}. Chalcogenides present as part of the material precursors are further screened based on their power factors (PFs) reported in the Materials Project⁴³. This results in the list of available precursors (**SI note 3** for the complete list).

As the first step, we measure the experimental PF for all the chalcogenide compositions in their original phase (i.e., without any alloying or doping). The error-correcting dense neural network (DNN) is then trained using the MagPie representation and predictions of *prior model* as inputs. The model is trained to predict the binary comparison variable using the experimentally measured in-lab power factor (refer to **equation (1)** for the definition of this binary variable). We optimize the hyperparameters and estimate the generalization using 5-fold cross-validation accuracy. The most generalizable model is then used to explore a new set of compositions for peak performance at 300K that is fed back, and the cycle is repeated. We explored different material engineering strategies in every round: composites, alloys and dopants (**SI note 4** for more details). We also list each of the compounds tested within every round of experimentation in **SI note 4**. With every round of experimentation, the ability of the error-correction model to account for the experimental changes improves (**Figure 1c** for the schematic on workflow). Multiple rounds of the feedback loop and exploration finally yield (PbSe):(2 wt% SnSb) as the optimal composition (refer to the section below for details on experimental characterization). We terminate the feedback loop if the top predicted material candidate remains unchanged after retraining the error-correction model using all the experimental data, including the latest round of experiments.

The results of hyperparameter optimizations are summarized in **Figure S4**. To establish a baseline, we also evaluate random forest classifiers and gradient boosting classifiers for their error-correction capacity using cross-validation accuracy (**SI note 5**) and find that they underperform as compared to DNNs.

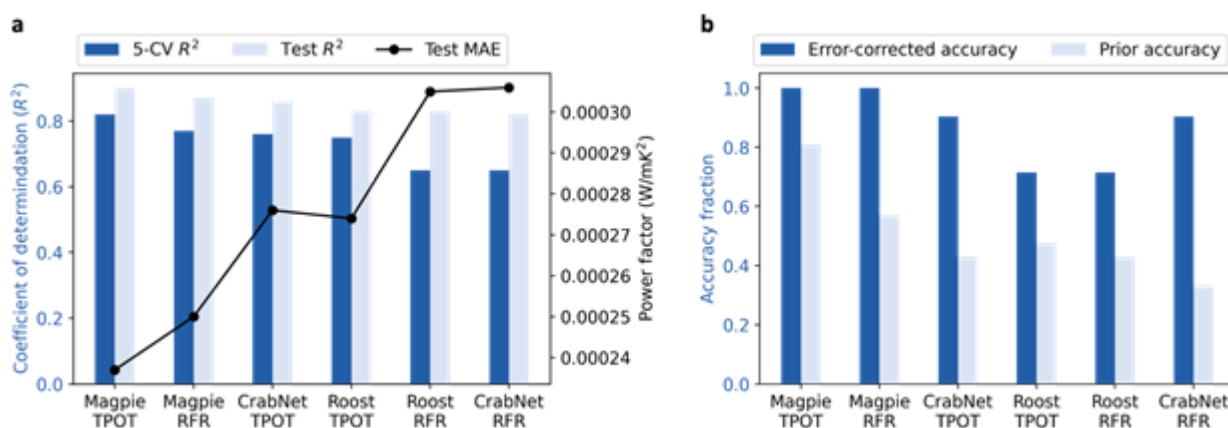


Figure 2. Performance of machine learning models with and without ECL on available and acquired data (a) Performance of the model with the best set of hyperparameters chosen based on the cross-validation performance quantified using mean absolute error (MAE) and coefficient of determination (R^2). We found that Magpie⁴⁰ based featurization outperformed transfer learning using Roost³⁸ and CrabNet³⁷. Here, the suffixes TPOT and RFR refer to the model used to belong to the one found by the TPOT library and Random Forest Regressor respectively. The ML model found by the TPOT library for all 3 representations is a combination of multiple ML models. Refer to **SI note 2** for the exact architecture and hyperparameters. **(b)** Performance for each model before and after performing ECL as described in the study. The prior accuracy is measured across all compositions and temperatures based on power factor ranking. The error-corrected accuracy is reported based on ordering at 219°C. ECL improves the accuracy for all cases. The improvement, however, depends upon both on featurization used and the initial prior model. The improvement varies as much as 0.16-0.40.

We compare the effect of change in the *prior model* on the accuracy of the error-correction learning (ECL) model (**Figure 2b**). For this, we train the error correction model on all the data collected from the experiments conducted before the last round of experiments. The accuracy is then evaluated on the last unseen round of experiments. We find that higher accuracy prior models lead to higher accuracy on error-corrected predictions. Interestingly, models sharing the same chemical representation also tend to perform similarly i.e., Magpie-based models, even though differing in prior model's accuracy show high error-corrected accuracies. Similarly, both the CrabNet-based models show similar accuracies post error correction. This observation aligns with the discussion before on the origin of higher accuracy observed with Magpie-based representation (**Figure S3**). This indicates that choosing an appropriate representation is essential while driving closed-loop experimentation. We also compared these results to the case where we used a weight-sharing network instead of a DNN. The trends were similar but with lower accuracy (**Figure S5**).

The effectiveness of error-correction learning depends upon the specific chemical representation utilized, a finding that highlights the sensitivity of the approach to the choice of representation. It therefore becomes important for practitioners to choose the appropriate chemical representation. The choice, as demonstrated in this study, can be made based on cross-validation accuracy once the first round of experiments has been performed. This choice can also be seen in the form of a trade-off between computation power and number of experiments. Validating and testing different representations to find better error-correcting models require more computation time (for instance a few GPU hours if training neural networks). This increase in computation cost allows us to reduce the number of experiments that need to be performed.

We find that the next-best model CrabNet, after error correction on the ranking of compounds, underperforms. As a result, at least one more round of experimentation is needed to reach the optimal candidate (Figure 2b). Thus, the accurate prior model developed herein reduces the number of experiments by at least $(1/5 \times 100\%) = 20\%$ (the total number of rounds of experiments in such a scenario would be 5). The best candidate composition $(\text{PbSe})_{0.98}(\text{SnSb})_{0.02}$ is ranked at 98th if the material candidates are ranked according to the order of decreasing power factors using the most accurate *prior model*. Thus, compared to the conventional high-throughput screening approach of testing the high-ranked predictions from the prior model, the current approach reduces the number of candidates that need to be tested by as much as 83% i.e., less than 1/3rd the number of experiments that would need to be conducted. We also compare our approach to a modified version of the previously proposed data fusion approach²⁸ in **Table 1** (refer to **SI note 6** for implementation details).

Table 1: We tabulate the accuracy and compare a former approach that requires defining rules for data fusion to our approach that utilizes dense neural networks for error correction.

Method	Accuracy on the final-round data at 219°C
Probabilistic constraint data-fusion ²⁸	0.57
ECL (this study)	0.99

2.3. Experimental validation

While the synthesis of inorganic compounds by solution method is tedious due to the limited number of suitable metal salts and solubility issues, vacuum techniques require an extensive amount of time for optimization. As such, we adopted ball milling to synthesize some of the predicted compounds owing to its high yield and faster optimization rate. We tested the top candidates from ML predictions and evaluated the compounds' thermoelectric properties. The measured properties were fed to the error correction module to improve the accuracy.

We observe that the highest electrical conductivity (σ) was obtained for SnSb with a low Seebeck coefficient (S). Stoichiometry change and forming composite with PbI₂ did not improve Seebeck of SnSb. However, its combination with PbSe resulted in balanced σ and S . Temperature-dependent σ and S (refer to the **characterization subsection of Methods** for more details) of the stoichiometric compounds, SnSe, WSe₂, GaTe and CrZnTe₄ exhibit semiconducting behavior while SnSb, PbSe, and NbSe₂ show metallic behavior. Notably, CrZnS₄ has extremely low electrical conductivity and consequently high Seebeck (**Figure 3a-c**). The highest electrical conductivity is observed in the case of SnSb (> 10000 S/cm) which leads to a very low Seebeck coefficient. PbSe exhibits the largest power factor among the stoichiometric compounds due to balanced electrical conductivity and Seebeck. Despite of extraordinary conductivity of SnSb, its modest power factor is a consequence of low Seebeck. In contrast, PbSe shows good Seebeck, but its electrical conductivity is not very high (We compare conductivities of many composites made and tested but not added to Figure 3b in **Figure S6**).

These findings align with the prediction that modest of SnSb to PbSe improves the power factor. We observe that small amounts of SnSb addition enhance the electrical conductivity of PbSe with a minor decrease in Seebeck (refer to **materials subsection** of Methods for details on synthesis and materials). Therefore, the PbSe/SnSb composition results in a high power factor of more than $850 \mu\text{W}/\text{mK}^2$ at room temperature for 2 wt% SnSb, and good performance is observed in the whole temperature regime (**Figure 3d-f**). Such a high-power factor for the present composite is comparable to state-of-the-art polycrystalline thermoelectric materials prepared by similar methods (**Figure S7**). High performance room-temperature thermoelectric materials are particularly lucrative for various applications including sensors, heat spreaders, internet of things and portable electronics⁴⁴. We would like to point out that most high-performance thermoelectric materials are synthesized by high-

temperature techniques in the form of single crystals whereas in the present work, polycrystalline materials are used without any extensive processing giving comparable performances at 300K. It should be noted that further performance improvement can be anticipated by additional processing techniques, such as spark plasma sintering, generally employed in the case of high-performance thermoelectric materials.

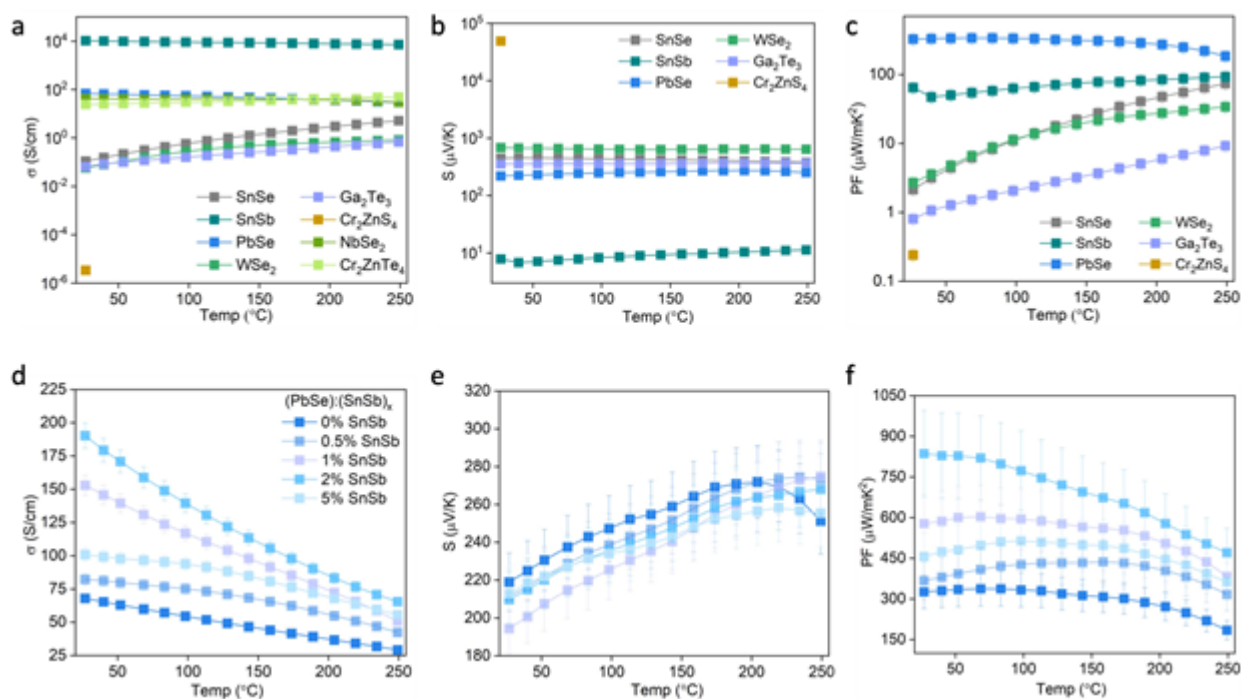


Figure 3: Thermoelectric performance of materials measured experimentally (a, b, c) show conductivity (σ), Seebeck coefficient (S), and Power Factor (PF) trends for stoichiometric compounds as a function of temperature. PbSe shows the highest power factor due to a balance between conductivity and Seebeck coefficient. (d, e, f) show σ , S and PF trends for our predicted most promising compositions $(\text{PbSe}):(\text{SnSb})_x$. We observe that the best composition has a power factor twice that of the undoped PbSe.

2.4. Origins of high power factor

To increase understanding of the atomistic origin of the high performance of $(\text{PbSe})_{0.98}(\text{SnSb})_{0.02}$, we perform Density Functional Theory (DFT) calculations (refer to the **DFT subsection of Methods** for details). We obtain the unit cell for the PbSe supercell from Materials Project⁴³ and use it to construct a 216 atoms supercell. Experimental X-ray diffraction (XRD) indicates that a small amount of alloying does not change the structure (**Figure 4a**). We use this information to optimize the geometry of the SnSb doped PbSe supercell by relaxing only the atomic positions and keeping the lattice constants fixed using

the Generalized Gradient Approximation (GGA) exchange-correlation (xc) functional (refer to **Figures S6-S7** for pictorial representation). We perform self-consistent HSE06 xc-functional calculations on the pristine and doped relaxed structures while incorporating spin-orbit coupling (SOC). Comparing the electronic structure of PbSe with PbSe:SnSb, we observe that the introduction of Sn and Sb modulates the density of states (DOS) near conduction band minimum (CBM) and valence band maximum (VBM) respectively. This modulation decreases the hole effective mass, improving the overall transport property of the PbSe:SnSb composite (**Figure 4b,c**) thereby resulting in a larger power factor. Please refer to **SI Figure S10-S11** for effective mass fitting plots.

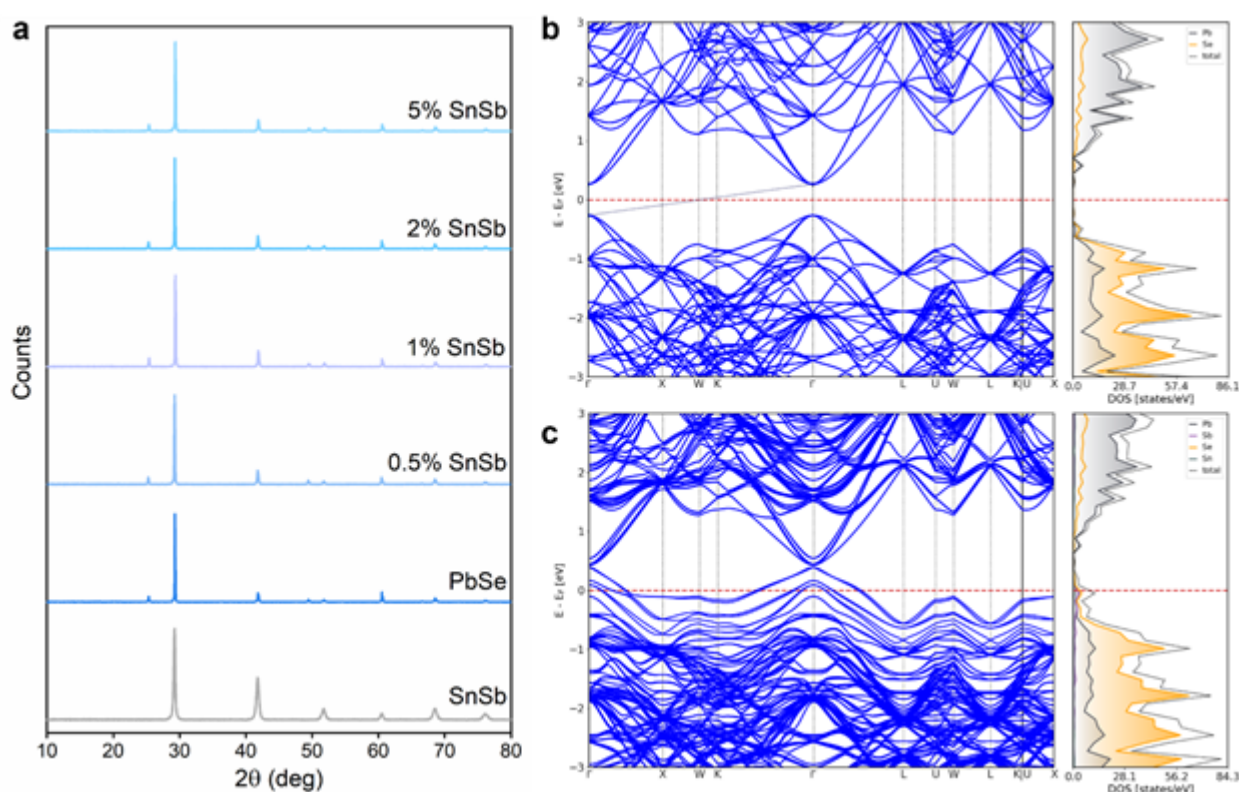


Figure 4: Structural and DFT analysis of the best thermoelectric candidates (a) shows XRD plots for $(\text{PbSe}):(\text{SnSb})_x$ composite. (b) shows the band structure of PbSe along the high-symmetry path. (c) shows band structure of $(\text{PbSe}):(2 \text{ wt}\% \text{ SnSb})$ along the high-symmetry path in the reciprocal space. The change in curvature of the state just below the defect state in the valence band leads to a reduction in the effective mass and therefore, improved transport properties (refer to **Figure S10-S11** for close-up).

3. Conclusions

In this study, we developed and demonstrated a two-step error correction learning approach to performing theory-driven closed-loop experimental exploration of thermoelectric

materials. In the process, we trained an ML model that outperformed prior models in the accurate prediction of experimental power factors. This improvement in accuracy reduces the number of in-lab experiments. The approach enables discovery of a thermoelectric material $(\text{PbSe})_{0.98}:(\text{SnSb})_{0.02}$ that exhibits a large power factor. DFT simulations show that the performance improvements are mainly a result of the chemical interactions and not structure distortions. The approach to thermoelectric materials may well be extended to the discovery of materials for applications requiring high conductivity and improved optoelectronic properties.

4. Experimental Section

Data:

We use the dataset collected by DopNet⁸ authors from the Materials Research Laboratory (MRL) for training the *prior models*. It is available at <https://github.com/ngs00/DopNet> and the original data can be found at <http://www.mrl.ucsb.edu:8080/datamine/thermoelectric.jsp>. Data based on our in-lab experiments is available at <https://github.com/hitarth64/TherML>.

Code availability

All the code and data used for this study are available as open source at <https://github.com/hitarth64/TherML>. We also provide all our trained models so that they can be used for future explorations.

Materials

Sn, Se, S, Zn, Te, and Cr powders were purchased from Sigma-Aldrich. PbSe, Sb, and WSe₂ were purchased from Alfa-Aesar. NbSe₂, GaSe, and Ga₂Te₃ were procured from American Elements.

SnSe, SnSb, Cr₂ZnS₄, and Cr₂ZnTe₄ were synthesized by ball milling stoichiometric amounts of precursors for 8, 15, 32, and 32 hours at 30 Hz, respectively. All composites were obtained by ball milling the respective compounds (e.g. SnSe and SnSb) for 20 minutes at 30 Hz.

Characterization

Electrical conductivity and Seebeck measurements were performed on Netzsch SBA 548 Nemesis thermoelectric set up under He environment. Samples were prepared by loading the powders into a steel compaction die and compressing them to form pellets for thermoelectric measurements. All pellets were annealed at 250 °C inside an N₂ glovebox before thermoelectric measurements. The instrument uncertainty for electrical conductivity and Seebeck measurements are ±5% and ±7%, respectively. As a result of the very low electrical conductivity of Cr₂ZnS₄, room temperature Seebeck was measured using a manual setup consisting of Peltier devices and a thermocouple. For thermal voltage measurements, a homemade setup was used with Peltier devices and thermocouples to apply the temperature gradient across the sample, and the voltage was recorded using a Keithley 6517B electrometer. XRD was measured using Bruker D8 Advance.

DFT

All DFT relaxations are performed using Vienna Ab initio Simulation Package (VASP) 6.2.1. We use Perdew-Burke-Ernzerhof (PBE)⁴⁵ exchange-correlation functional for geometry optimization of the material systems. The energy cut-off for the plane wave was set to 520 eV. The energy threshold for self-consistent energy convergence was set to 10⁻³ eV whereas the global convergence threshold was set to 10⁻² eV energy difference between two successive ionic steps. A smearing width of 0.05 eV, consistent with the Materials Project, was chosen for calculations. Calculations were performed over a uniform k-points grid of 2 × 2 × 2 generated using the Monkhorst scheme⁴⁶.

Given the optimized geometry, we performed static DFT calculation using hybrid HSE06⁴⁷ functional while incorporating the Spin-orbit coupling (SOC) effect over high-symmetry points of the reciprocal lattice using all-electron DFT code FHI-AIMS.

Supporting Information

Supporting Information is available from the Wiley Online Library or the author.

Acknowledgments

M.A.H. and H.C. contributed equally to this work. This publication is supported by the King Abdullah University of Science and Technology (KAUST) Office of Sponsored Research (OSR) under Award No. OSR-CRG2018-3737. TOC was created by Ana Bigio, scientific illustrator at KAUST. ML models were trained using QUEST computing clusters located at

Northwestern University. DFT calculations were performed both at the QUEST computing cluster located at Northwestern University and the Narval computing cluster which is part of Compute Canada and made accessible through University of Toronto. We thank Lydia Li for help in designing Figure 1.

Received: ((will be filled in by the editorial staff))

Revised: ((will be filled in by the editorial staff))

Published online: ((will be filled in by the editorial staff))

References

1. Jood, P., Ohta, M., Yamamoto, A. & Kanatzidis, M. G. Excessively Doped PbTe with Ge-Induced Nanostructures Enables High-Efficiency Thermoelectric Modules. *Joule* **2**, 1339–1355 (2018).
2. Tan, G., Ohta, M. & Kanatzidis, M. G. Thermoelectric power generation: from new materials to devices. *Philosophical Transactions of the Royal Society A* **377**, (2019).
3. Vostrikov, S., Somov, A. & Gotovtsev, P. Low temperature gradient thermoelectric generator: Modelling and experimental verification. *Appl Energy* **255**, 113786 (2019).
4. van Toan, N., Thi Kim Tuoi, T., van Hieu, N. & Ono, T. Thermoelectric generator with a high integration density for portable and wearable self-powered electronic devices. *Energy Convers Manag* **245**, 114571 (2021).
5. Davies, D. W. *et al.* Computational Screening of All Stoichiometric Inorganic Materials. *Chem* **1**, 617–627 (2016).
6. Choubisa, H. *et al.* Crystal Site Feature Embedding Enables Exploration of Large Chemical Spaces. *Matter* (2020) doi:10.1016/j.matt.2020.04.016.
7. Choubisa, H. *et al.* Accelerated chemical space search using a quantum-inspired cluster expansion approach. (2022) doi:10.48550/arxiv.2205.09007.
8. Na, G. S., Jang, S. & Chang, H. Predicting thermoelectric properties from chemical formula with explicitly identifying dopant effects. *npj Computational Materials* **2021 7:1 7**, 1–11 (2021).
9. Chakraborty, S. *et al.* Rational Design: A High-Throughput Computational Screening and Experimental Validation Methodology for Lead-Free and Emergent Hybrid Perovskites. *ACS Energy Lett* **2**, 837–845 (2017).
10. Ding, G., Gao, G. & Yao, K. High-efficient thermoelectric materials: The case of orthorhombic IV-VI compounds. *Scientific Reports* **2015 5:1 5**, 1–7 (2015).
11. Jain, A., Shin, Y. & Persson, K. A. Computational predictions of energy materials using density functional theory. *Nature Reviews Materials* **2016 1:1 1**, 1–13 (2016).

12. Zhu, H. *et al.* Computational and experimental investigation of TmAgTe₂ and XYZ₂ compounds, a new group of thermoelectric materials identified by first-principles high-throughput screening. *J Mater Chem C Mater* **3**, 10554–10565 (2015).
13. Hodges, J. M. *et al.* Two-Dimensional CsAg₅Te₃- xS_x Semiconductors: Multi-anion Chalcogenides with Dynamic Disorder and Ultralow Thermal Conductivity. *Chemistry of Materials* **30**, 7245–7254 (2018).
14. Vikram, Sahni, B., Jain, A. & Alam, A. Quasi-2D Carrier Transport in KMgBi for Promising Thermoelectric Performance. *ACS Appl Energy Mater* **5**, 9141–9148 (2022).
15. Barreteau, C., Crivello, J. C., Joubert, J. M. & Alleno, E. Looking for new thermoelectric materials among TMX intermetallics using high-throughput calculations. *Comput Mater Sci* **156**, 96–103 (2019).
16. Jain, A. & McGaughey, A. J. H. Thermal transport by phonons and electrons in aluminum, silver, and gold from first principles. *Phys Rev B* **93**, 081206 (2016).
17. Andrejevic, N. *et al.* Machine-Learning Spectral Indicators of Topology. *Advanced Materials* **34**, 2204113 (2022).
18. Tshitoyan, V. *et al.* Unsupervised word embeddings capture latent knowledge from materials science literature. *Nature* **571**, 95–98 (2019).
19. Pal, K., Park, C. W., Xia, Y., Shen, J. & Wolverton, C. Scale-invariant machine-learning model accelerates the discovery of quaternary chalcogenides with ultralow lattice thermal conductivity. *npj Computational Materials* **2022 8:1** **8**, 1–12 (2022).
20. Ren, Z. *et al.* An invertible crystallographic representation for general inverse design of inorganic crystals with targeted properties. *Matter* **5**, 314–335 (2022).
21. Wang, N. *et al.* Machine-Learning-Accelerated Development of Efficient Mixed Protonic–Electronic Conducting Oxides as the Air Electrodes for Protonic Ceramic Cells. *Advanced Materials* **34**, 2203446 (2022).
22. Pohls, J. H. *et al.* Experimental validation of high thermoelectric performance in RECuZnP₂ predicted by high-throughput DFT calculations. *Mater Horiz* **8**, 209–215 (2021).
23. Miyata, M. *et al.* High-Throughput Screening of Sulfide Thermoelectric Materials Using Electron Transport Calculations with OpenMX and BoltzTraP. *J Electron Mater* **47**, 3254–3259 (2018).
24. Jia, T., Feng, Z., Guo, S., Zhang, X. & Zhang, Y. Screening Promising Thermoelectric Materials in Binary Chalcogenides through High-Throughput Computations. *ACS Appl Mater Interfaces* **12**, 11852–11864 (2020).
25. Choudhary, K., Garrity, K. F. & Tavazza, F. Data-driven discovery of 3D and 2D thermoelectric materials. *Journal of Physics: Condensed Matter* **32**, 475501 (2020).

26. Hu, Z., Wu, W., Wang, Q. & Shao, X. Theoretical Study of Ba₂X₆ (X = S, Se, Te) for Thermoelectric Applications Based on First-Principles Calculations and Machine Learning. *Journal of Physical Chemistry C* **126**, 12735–12741 (2022).
27. Hou, Z., Takagiwa, Y., Shinohara, Y., Xu, Y. & Tsuda, K. Machine-Learning-Assisted Development and Theoretical Consideration for the Al₂Fe₃Si₃ Thermoelectric Material. *ACS Appl Mater Interfaces* **11**, 11545–11554 (2019).
28. Sun, S. *et al.* A data fusion approach to optimize compositional stability of halide perovskites. *Matter* **4**, 1305–1322 (2021).
29. Wagner, J. *et al.* The evolution of Materials Acceleration Platforms: toward the laboratory of the future with AMANDA. *J Mater Sci* **56**, 16422–16446 (2021).
30. MacLeod, B. P. *et al.* Self-driving laboratory for accelerated discovery of thin-film materials. *Sci Adv* **6**, (2020).
31. Burger, B. *et al.* A mobile robotic chemist. *Nature* **2020 583:7815** **583**, 237–241 (2020).
32. Kusne, A. G. *et al.* On-the-fly closed-loop materials discovery via Bayesian active learning. *Nature Communications* **2020 11:1** **11**, 1–11 (2020).
33. Wu, M. *et al.* Target-Driven Design of Deep-Ultraviolet Nonlinear Optical Materials via Interpretable Machine Learning. *Advanced Materials* 2300848 (2023) doi:10.1002/ADMA.202300848.
34. Lampe, C. *et al.* Rapid Data-Efficient Optimization of Perovskite Nanocrystal Syntheses through Machine Learning Algorithm Fusion. *Advanced Materials* 2208772 (2023) doi:10.1002/ADMA.202208772.
35. Park, C. W. & Wolverton, C. Developing an improved Crystal Graph Convolutional Neural Network framework for accelerated materials discovery. *Phys Rev Mater* **4**, (2019).
36. Xie, T. & Grossman, J. C. Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties. *Phys Rev Lett* **120**, 145301 (2018).
37. Wang, A. Y. T., Kauwe, S. K., Murdock, R. J. & Sparks, T. D. Compositionally restricted attention-based network for materials property predictions. *npj Computational Materials* **2021 7:1** **7**, 1–10 (2021).
38. Goodall, R. E. A. & Lee, A. A. Predicting materials properties without crystal structure: deep representation learning from stoichiometry. *Nature Communications* **2020 11:1** **11**, 1–9 (2020).
39. Askerka, M. *et al.* Learning-in-Templates Enables Accelerated Discovery and Synthesis of New Stable Double Perovskites. *J Am Chem Soc* **141**, 3682–3690 (2019).
40. Ward, L., Agrawal, A., Choudhary, A. & Wolverton, C. A general-purpose machine learning framework for predicting properties of inorganic materials. *npj Computational Materials* **2016 2:1** **2**, 1–7 (2016).

41. Gaultois, M. W. *et al.* Data-driven review of thermoelectric materials: Performance and resource considerations. *Chemistry of Materials* **25**, 2911–2920 (2013).
42. Le, T. T., Fu, W. & Moore, J. H. Scaling tree-based automated machine learning to biomedical big data with a feature set selector. *Bioinformatics* **36**, 250–256 (2020).
43. Jain, A. *et al.* Commentary: The materials project: A materials genome approach to accelerating materials innovation. *APL Materials* vol. 1 011002 Preprint at <https://doi.org/10.1063/1.4812323> (2013).
44. Chen, J. *et al.* Ternary Ag₂Se_{1-x}Tex: A Near-Room-Temperature Thermoelectric Material with a Potentially High Figure of Merit. *Inorg Chem* **60**, 14165–14173 (2021).
45. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys Rev Lett* **77**, 3865–3868 (1996).
46. Monkhorst, H. J. & Pack, J. D. Special points for Brillouin-zone integrations. *Phys Rev B* **13**, 5188 (1976).
47. Heyd, J. & Scuseria, G. E. Efficient hybrid density functional calculations in solids: Assessment of the Heyd–Scuseria–Ernzerhof screened Coulomb hybrid functional. *J Chem Phys* **121**, 1187 (2004).

Table of Contents text:

A machine learning model using error-correction learning with experimental feedback is developed to predict high performance thermoelectric materials. Experimentally a new family of thermoelectric materials consisting of PbSe:SnSb is achieved along high room temperature power factor.

H. Choubisa,[†] M. A. Haque,[†] T. Zhu, L. Zeng, M. Vafaie, D. Baran,^{*} E. H. Sargent^{*}

Closed-loop Error Correction Learning Accelerates Experimental Discovery of Thermoelectric Materials

Table of Contents text:

Table of Contents Image:

