

Automatic exhaustive calculations of large material space by Korringa-Kohn-Rostoker coherent potential approximation method applied to equiatomic quaternary high entropy alloys

T. Fukushima^{1,2,3,4,*}, H. Akai¹, T. Chikyow,⁵ and H. Kino⁵


¹*Institute for Solid State Physics, The University of Tokyo, Kashiwa, Chiba 277-8581, Japan*

²*Center for Spintronics Research Network, Osaka University, Toyonaka, Osaka 560-8531, Japan*

³*Center for Spintronics Research Network, The University of Tokyo, Bunkyo-ku, Tokyo 113-8656, Japan*

⁴*Institute for AI and Beyond, The University of Tokyo, Bunkyo-ku, Tokyo 113-8656, Japan*

⁵*MaDIS, National Institute for Materials Science, Tsukuba, Ibaraki 305-0044, Japan*

 (Received 30 July 2021; revised 2 November 2021; accepted 10 December 2021; published 17 February 2022)

Automatic exhaustive exploration of a large material space by high-performance supercomputers is crucial for developing new functional materials. We demonstrated the efficiency of high-throughput calculations using the all-electron Korringa-Kohn-Rostoker coherent potential approximation method with the density functional theory for the large material space consisting of quaternary high entropy alloys, which are nonstoichiometric and substitutionally disordered materials. The exhaustive calculations were performed for 147 630 systems based on the AkaiKKR program package and supercomputer Fugaku, where the numerical parameters and self-consistent convergence are automatically controlled. The large material database including the total energies, magnetization, Curie temperature, and residual resistivity was constructed by our calculations. We used frequent itemset mining to identify the characteristics of parcels in magnetization and Curie temperature space. We also identified the elements that enhance the magnetization and Curie temperature and clarified the rough dependence of the elements through regression modeling of the residual resistivity.

DOI: [10.1103/PhysRevMaterials.6.023802](https://doi.org/10.1103/PhysRevMaterials.6.023802)

I. INTRODUCTION

First-principles calculations, which can accurately investigate the electronic structures, magnetic properties, and transport properties of materials without any empirical parameters, are effective tools for designing new functional materials. Although there are limitations in the unit cell sizes and the approximations for the exchange-correlation potentials, they have been applied to several systems, such as cathode materials [1], photovoltaic absorbers [2], and spintronics materials [3–5], and their effectiveness has been confirmed. Recently, with the development of computer hardware and numerical algorithms, material databases have been actively constructed from high-throughput calculations, i.e., exhaustive calculations, based on first-principles approaches. By applying machine learning techniques to the constructed material databases, one can elucidate the mechanisms behind the physical and chemical properties in the target and accelerate the discovery of new materials. The databases containing a wide variety of material data are required to efficiently perform machine learning. However, performing exhaustive calculations with high speed and high accuracy for such large material space is not easy. The validity of material-dependent numerical parameters should always be considered. For ex-

ample, cut-off parameters, potential mixing parameters, and the selection of pseudopotentials affect calculation results and convergence. We state two important criteria for constructing material databases by automatic exhaustive calculations: (i) the understanding of underlying physics and chemistry in target material space and (ii) the knowledge of the first-principles calculation program package, that one uses. Automatic exhaustive calculations, which do not satisfy the above two criteria might lead to unreliable material databases.

There are several material databases, such as the novel materials discovery (NOMAD) [6], the open quantum materials database (OQMD) [7,8], the automatic flow for materials discovery (AFLOW) [9], and the materials project (MP) [10]. Regardless of the data validity, researchers can freely access these databases and obtain useful material information. However, they mainly focus on the electronic structures and structural stabilities of simple substances and stoichiometric compounds. They contain no detailed data for magnetic properties (e.g., magnetic exchange interactions and magnetic transition temperatures) and transport properties (e.g., electric conductivity) in nonstoichiometric systems which are essential for developing magnetic and spintronic materials. This is because it is necessary to use a scale-bridging simulation technique [4] and a linear response approach [11] with more computational costs and a complicated numerical algorithm for evaluating the above physical quantities, in addition to a simple search of an electronic ground state within the density functional theory (DFT) [12,13]. Another reason is the difficulty in handling the nonstoichiometric systems in DFT calculations. A supercell method is usually used for treating the nonstoichiometric systems. This method can calculate atomic forces and perform structural optimizations for a

*fuku@issp.u-tokyo.ac.jp

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

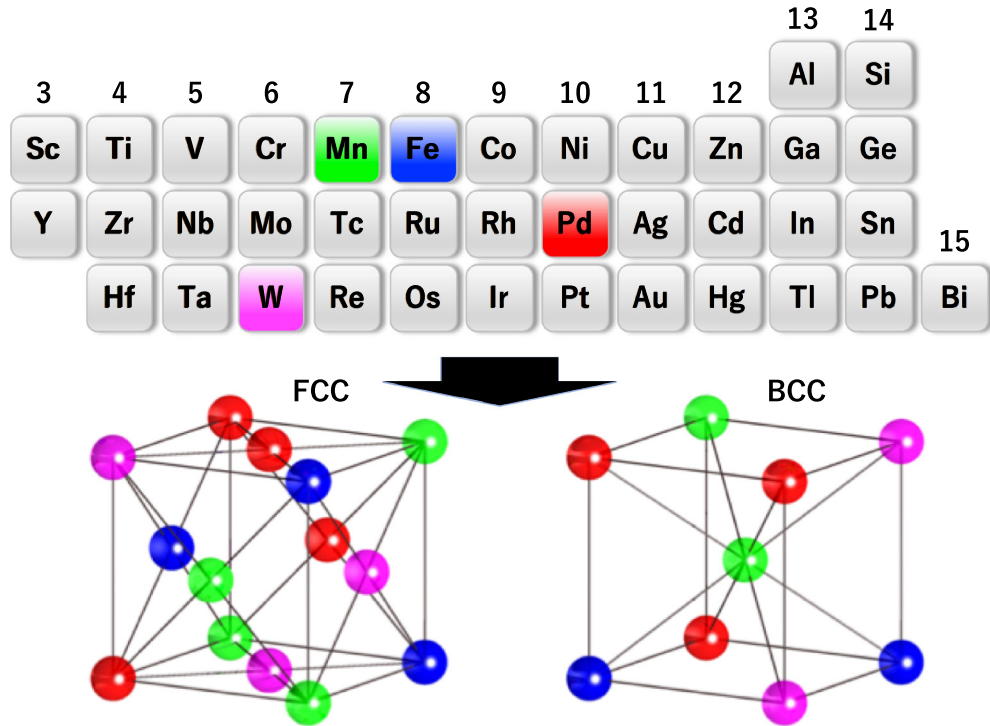


FIG. 1. Material space for the high-throughput calculations. The target elements are indicated in the upper part. The four elements are selected, and the BCC and FCC solid solution phases are created. The total number of systems in the material space is 147 630. In the figure, the case of MnFePdW is shown as an example.

specific configuration. However, one requires large supercells and huge computational costs to avoid the artificial Coulomb interaction originating from the periodicity and investigate the configurational dependence of the physical quantities.

In this study, we demonstrate the efficiency of high-throughput calculations for equiatomic quaternary high entropy alloys which are nonstoichiometric and configurational disordered systems, on the basis of the all-electron Korringa-Kohn-Rostoker (KKR) Green’s function method [14,15] with the coherent potential approximation (CPA) [16,17]. We employ the AkaiKKR program package developed by Akai [18,19]. The numerical parameters and convergence procedure are automatically controlled. We construct a large-scale material database, including the simple electronic structures and the finite temperature magnetic and the transport properties without any empirical parameters. The constructed material database is essential for screening new magnetic and spintronic materials. Based on this large-scale database, we also identify important elements and factors which determine the magnetic and transport properties of equiatomic quaternary high entropy alloys by data science approaches.

II. TARGET MATERIAL SPACE

As a target material space, we choose quaternary high entropy alloys with face-centered cubic (FCC) and base-centered cubic (BCC) solid solution phases, where the four principal elements have the same atomic concentration and are randomly distributed in the lattices. The number of target elements are 38 as follows: Al, Si, Sc, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, Ga, Ge, Y, Zr, Nb, Mo, Tc, Ru, Rh, Pd, Ag, Cd,

In, Sn, Hf, Ta, W, Re, Os, Ir, Pt, Au, Hg, Tl, Pb, and Bi. We select four elements from the above target elements and form an equiatomic quaternary high entropy alloy (Fig. 1). Since both BCC and FCC cases are considered, the total number of systems is 147 630. We perform exhaustive calculations for such large material space on the flagship-supercomputing system in Japan, Fugaku, developed by RIKEN Center for Computational Science.

III. CALCULATION METHODS

We conducted the electronic structure calculations using the AkaiKKR code [18], where the all-electron KKR Green’s function method [14,15] is implemented within DFT [12,13]. The KKR Green’s function method is based on the multiple scattering theory, and one electron Green’s function for a Kohn-Sham equation is directly calculated. The electron density of a system [$\rho(\mathbf{r})$] is obtained from the imaginary part of the site-diagonal retarded Green’s function integrated up to the Fermi level (E_F):

$$\rho(\mathbf{r}) = \sum_i |\phi_i(\mathbf{r})|^2 - \frac{1}{\pi} \lim_{r' \rightarrow r} \times \Im \int_{-\infty}^{E_F} (G^s(\mathbf{r}, \mathbf{r}', \epsilon) + G^m(\mathbf{r}, \mathbf{r}', \epsilon)) d\epsilon, \quad (1)$$

where $\phi_i(\mathbf{r})$ is the core state, $G^s(\mathbf{r}, \mathbf{r}', \epsilon)$ is the Green’s function for a single-site potential (i.e., potential part), and $G^m(\mathbf{r}, \mathbf{r}', \epsilon)$ contains the multiple scattering effects (i.e., back scattering part). As well as other all-electron calculation schemes, valence and core electrons are separately handled

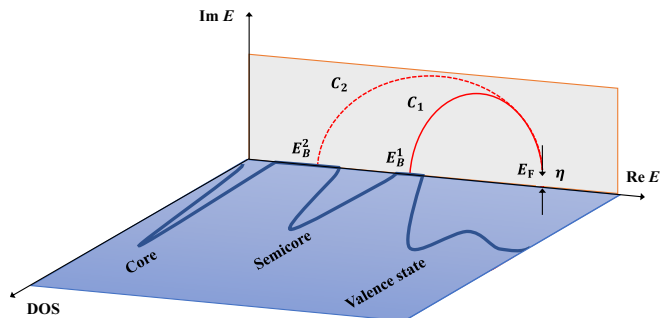


FIG. 2. Complex energy contour (C_1 or C_2) used in the AkaiKKR code. The energy contour ranges from the bottoms of the valence bands (E_B^1) or semicore states (E_B^2) to the Fermi level (E_F). To make it easier to understand, density of states (DOS) is schematically shown in the figure. η is a small imaginary part attached to E_F .

in actual calculations. The Green's function is defined for the valence electrons, and the core electrons are treated as in atomic-like fashion. Since the Green's function is strongly structured close to the real axis, one needs to many mesh points for accurate energy integration. To avoid such problems, the energy integral is evaluated on a semicircular contour in the complex energy plane in the AkaiKKR code, since Green's function is analytic in the upper half-plane of complex energy. This procedure decreases the number of energy mesh points and stable numerical simulations. To perform the energy integration in Eq. (1), we need to determine the interval of the integration, as shown in Fig. 2. The upper limit of the interval corresponds to E_F , and the lower limit (E_B) is generally set below the valence bands (C_1 in Fig. 2). When there exist shallow semicore states, such as Zn-3d and Ga-3d states, we should include these states in energy contour (C_2 in Fig. 2); otherwise, wrong electronic structures, magnetism, and transport properties will be obtained for alloy systems containing the above elements. A small imaginary part (η) is attached to E_F in actual calculations since Green's function cannot be defined due to a branch cut along a positive real axis; however such a problem does not occur at the bottom of the interval. This imaginary part yields smearing effects and results in numerical stability, particularly for metallic systems. Other smearing methods (e.g., Gaussian and Fermi-Dirac distribution functions) are not used in the AkaiKKR code. Typically, $\eta = 10^{-3}$ – 10^{-4} Ry is employed for self-consistent calculations. Although a smaller η produces a secure calculation result, many k -point meshes are required for convergence. To obtain the accurate electronic structures and good convergence properties in the high-throughput calculations, the energy width and η should be automatically determined.

The Curie temperatures are evaluated based on a scale-bridging method, where first-principles calculations of magnetic exchange interactions (J_{ij}) are coupled to the classical Heisenberg model. We employ Liechtenstein's formula with the magnetic force theorem for calculating J_{ij} [20]:

$$J_{ij} = \frac{1}{4\pi} \Im \int^{E_F} dE \text{Tr}_L \{ \Delta_i T_{\uparrow}^{ij} \Delta_j T_{\downarrow}^{ij} \}, \quad (2)$$

where $\Delta_i = t_{i\uparrow}^{-1} - t_{i\downarrow}^{-1}$, with $t_{i\uparrow(\downarrow)}$ being the single-site t -matrix at site i for the spin-up (down) state. These quantities

are evaluated using the ground state potential determined by self-consistent calculations. $T_{\uparrow(\downarrow)}^{ij}$ is the scattering path operator between sites i and j for the spin up (down) state. Tr_L is the trace over the orbital variables, i.e., azimuthal l and magnetic m quantum numbers. If the J_{ij} values are obtained, one can estimate T_C using the mean-field approximation, cluster-like approximation [21,22], Tyablicov approximation [23,24], and Monte Carlo simulation. The above approach has been applied to several disordered materials [4,25]. Here, we use the mean-field approximation. Although the mean-field approximation tends to overestimate the experimentally observed T_C , it is considered that the chemical trend can be acquired.

For the transport property of the target materials, we focus on the residual electric resistivity. Generally, frequency-dependent electric conductivity tensors are described by linear-response theory [26]. A more practical calculation scheme for the diagonal part of the conductivity tensor along the z direction (σ_{zz}) of an atomic disordered system is Kubo-Greenwood formula [27]. In terms of retarded Green's function ($G^+(\epsilon)$), σ_{zz} is written as

$$\sigma_{zz}(\epsilon) = \frac{\hbar}{\pi V} \text{Tr} \langle \hat{j}_z \Im G^+(\epsilon) \hat{j}_z \Re G^+(\epsilon) \rangle_{\text{conf}}. \quad (3)$$

Here, V is the volume of the unit cell, \hat{j}_z is the current operator, and $\langle \dots \rangle_{\text{conf}}$ means a configuration average of possible atomic disorders. Butler presented a prescription to calculate Eq. (3) in the framework of KKR Green's function formalism, and showed that $\langle \dots \rangle_{\text{conf}}$ can be treated by the CPA approach. The matrix elements of \hat{j}_z are calculated using on the energy shell solution of the single-site Schrödinger equation. We employ the same expression as Butler's. Readers can find the details of this formula in Refs. [11,28]. In our calculations, vertex corrections are included.

IV. CALCULATION CONDITIONS

The high-throughput calculations are performed under the condition that the atomic arrangements of the equiatomic quaternary high entropy alloys are completely random. The atomic disorders are treated using the CPA approach. In the CPA approach, the multiple scattering effects due to the impurity potentials are replaced by an effective medium described by coherent t -matrix [16,17]. CPA can efficiently calculate disordered systems without using a large supercell, leading to the significant reduction of computational costs. We notice here that since CPA is a single-site approach, the local environment effect, short-range atomic order, and structural optimization, are not considered. These effects lead to lattice distortion and sometimes affect the local electronic structure and magnetism in high entropy alloys. However, it also has to be noticed that Fukushima *et al.* and Ogura *et al.* showed that, for CrFeCoNi and $\text{Al}_x\text{CrFeCoNi}$ with random solid solution phase, the configurational averaged quantities by CPA are consistent with the results of the large-scale supercell calculations [29,30].

The lattice constants are estimated from the atomic volume of each constituent element observed in experiments [31]. The exchange-correlation parts are treated using the generalized gradient approximation parametrization of Perdew, Becke, and Erzenhof (PBE) [32]. Although the angular momentum

cutoff, $l_{\max} = 2$, gives reasonable results for the magnetic moments and T_C , the conductivity calculations are sensitive to l_{\max} . Therefore we use $l_{\max} = 3$ to include f -scattering effects [28]. The overlapping atomic sphere approximation (ASA) is used in the high-throughput calculations. When calculating atomic t -matrix, the same atomic sphere radius for each constituent element is employed. The space integral for the current operator in the Kubo-Greenwood formula is taken for this atomic sphere. The relativistic effects are included through the scalar relativistic approximation [33] and the spin-orbit interaction is not considered. It generally affects the electronic structure (e.g., the splitting of the core states) for the nonmagnetic systems containing heavy elements, such as Hg, Pb, and Bi. ASA calculations sometimes cause difficulties in convergence, when the shallow and localized $4f$ states in Hf and Ta atoms are included in the valence states. To avoid this difficulty, we treat these states as the core states, so that the resonances originating from the $4f$ states are eliminated [34].

First, by rough self-consistent calculations with lower k -sampling points, we search candidates of E_B (see Fig. 2), which is more than 0.1 Ry away from the core states, semi-core states, and the bottom of valence bands. Then, the tight self-consistent calculations with 256 k -sampling points in the first irreducible Brillouin zone are performed for one of the selected E_B . The number of mesh points along the complex energy integral path is fixed to 65. We carry out the self-consistent calculations with the small imaginary part attached to E_F , i.e., $\eta = 10^{-4}$ Ry; however, if the convergence problem is encountered, the condition, $\eta = 10^{-3}$ Ry, is used. For the BCC and FCC phases, the same interval of the energy integral and η are chosen. Starting with the initial potential and coherent t -matrix constructed from the superposition of the atomic potentials and t -matrices, we perform the self-consistent calculation of the ground state potential and coherent t -matrix. At the beginning of the calculations, where the charge neutrality oscillates and the root mean square (RMS) errors for the potential are large, we employ the Chebyshev iterative method [35]. When the RMS error in potential becomes small, the mixing scheme is switched to Broyden's second-mixing method to accelerate the convergence. During the iteration, with reference to the history of the charge neutrality, total magnetic moment, and RMS error, the E_B , η , and mixing parameter are appropriately adjusted. If the RMS error is smaller than 10^{-6} Ry, the self-consistent iteration is stopped. T_C and transport properties are calculated using Green's function determined by self-consistent calculation. In the transport calculations, 34,061 k -sampling points are used and η is set to 10^{-5} Ry. We execute high-throughput calculations automatically for the target materials space according to the above calculation procedure.

V. RESULTS

A. Notations

First, some notations are defined. We consider a randomly distributed solid solution, which is called high entropy alloys when the number of the constituent elements is many. Let $\mathcal{X}\mathcal{Y}\mathcal{Z}\mathcal{W}$ be the constituent elements for the equiatomic quaternary high entropy alloys (HEA4). We analyze the bulk and local properties of HEA4. Hereafter, M is the magnetiza-

TABLE I. The number of data instances in the BCC and FCC phases and in P_A (a set of all materials) and P_M (a set of all magnetic materials) and P_N (a set of all nonmagnetic materials) defined in Eqs. (4)–(6).

	P_A	P_M	P_N
BCC	73,145	21,900	51,245
FCC	73,143	21,205	51,938

tion in tesla (T), T_C is the magnetic transition temperature in Kelvin (K), R is the residual resistivity in $\mu\Omega$ cm, and $m_{\mathcal{X}}$ is the magnetic moment for the element \mathcal{X} in μ_B . We use thresholds of the local magnetic moments defined by $m_{\text{th},0} = 0.001 \mu_B$ and $m_{\text{th},1} = 0.2 \mu_B$. The former is used to identify magnetic/nonmagnetic materials explained below, the latter is used to define the spin configuration later.

Next, several sets of elements and their combinations are defined. Sets of 38 elements, 3d magnetic elements, and other elements are defined as follows:

$$E_A := \{\text{Al, Si, Sc} - \text{Ge, Y} - \text{Sn, Hf} - \text{Bi}\},$$

$$E_M := \{\text{Cr, Mn, Fe, Co, Ni}\},$$

$$E_N := E_A/E_M.$$

Sets of elemental combinations for all/magnetic/nonmagnetic materials are defined as follows:

$$P_A := \{\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W} \mid \mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W} \in \mathcal{P}_4(E_A)\}, \quad (4)$$

$$P_M := \{\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W} \mid \mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W} \in \mathcal{P}_4(E_M), |m(\mathcal{X})| > m_{\text{th},0}, \dots, |m(\mathcal{W})| > m_{\text{th},0}\}, \quad (5)$$

$$P_N := P_A/P_M, \quad (6)$$

where a set formed by all k -combinations made of a set E without any duplicates is denoted by $\mathcal{P}_k(E)$. Notably, P_A , P_M , and P_N in the BCC phases are different from those in the FCC phases.

B. Overview of calculated data

Automatic exhaustive calculations were performed on Fugaku by parallelization using 2000 nodes, where the material space was divided into each node. Our high-throughput calculation scheme showed that approximately 99.1% of the material space, i.e., 146 288 of 147 630 systems, satisfied the convergence criterion after data cleansing for apparently strange Ga local magnetic moments and for AlMnZnBi, which had quite large T_C , but its dependence on energy width was large. Detailed numbers are listed in Table I. The calculated data are placed in Zenodo repository [36].

C. Overall distribution of physical properties

The distributions of total magnetization per volume (M), magnetic phase transition temperature (T_C), and residual resistivity (R) for P_A in the BCC phase is shown in Fig. 3. (a) shows the distribution of physical properties of P_A . A large fraction of the total magnetization is zero, which belongs to P_N . Meanwhile, the distribution and pair plot of only magnetic

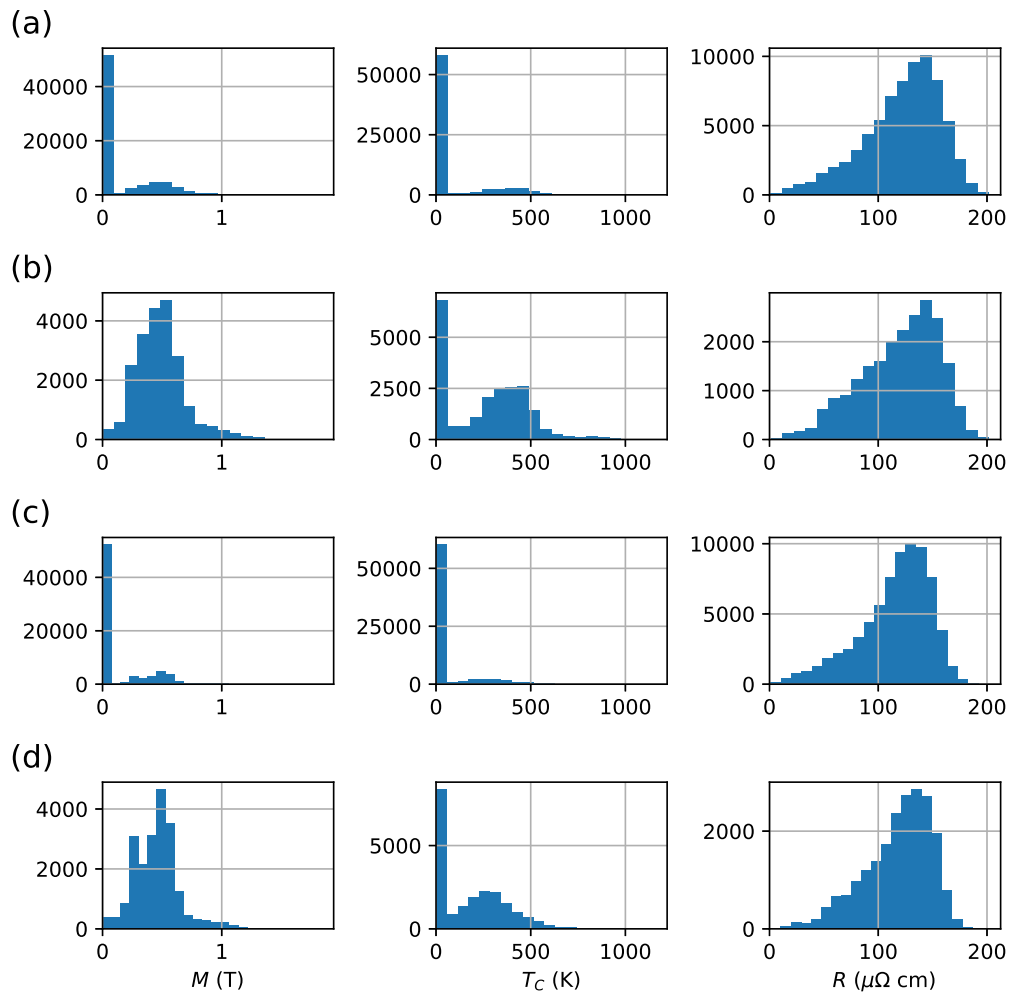


FIG. 3. Distribution and pair plot of M , T_C and R of HEA4. (a) shows the histogram for P_A (a set of all materials) and (b) for P_M (a set of all magnetic materials) in the BCC phase. (c) shows the histogram for P_A , and (d) for P_M in the FCC phase. P_A and P_M are defined in Eqs. (4) and (5).

materials, P_M , are shown in (b). The contribution of almost zero total magnetization reduced, but the proportion of almost zero magnetic T_C was still large. By contrast, the shape of the distribution of R did not change significantly between (a) and (b), which was partly because the change in R between P_M and P_N was small relative to the overall scale, although the former had a larger R . The FCC phase had almost the same characteristics and is shown in Figs. 3(c) and 3(d).

D. Magnetic properties

We show the characteristic features of magnetic properties; Fig. 4 shows the scatter plots as a function of the magnetization and T_C for the (a) BCC and (b) FCC phases with markers showing $\text{FeCo}\mathcal{X}\mathcal{Y}$, $\text{FeCoNi}\mathcal{Y}$, $\text{MnFeCo}\mathcal{Y}$, $\text{MnFeNi}\mathcal{Y}$, and $\text{MnFe}\mathcal{X}\mathcal{Y}$, which have high M s and high T_C s. The other materials are indicated by the gray points in the figure; \mathcal{X} and \mathcal{Y} are elements in E_N . Their number of instances was, respectively, 625, 618, 35, 35, and 35 in the BCC phase and 529, 512, 20, 26, and 21 in the FCC phase in P_M .

As $\mathcal{X}\mathcal{Y} \in P_A$, in the BCC phase, $\text{FeCo}\mathcal{X}\mathcal{Y}$ often has a high T_C of more than 700 K. $\text{MnFe}\mathcal{X}\mathcal{Y}$ had a moderate T_C but a larger M than $\text{FeCo}\mathcal{X}\mathcal{Y}$, and the boundary between $\text{FeCo}\mathcal{X}\mathcal{Y}$

and $\text{MnFe}\mathcal{X}\mathcal{Y}$ often exists with respect to T_C . $\text{MnFeCo}\mathcal{X}$ had the largest M and the largest T_C . $\text{FeCoNi}\mathcal{X}$ had the largest T_C and large M . Meanwhile, in the FCC phase, separation of $\text{FeCo}\mathcal{X}\mathcal{Y}$ and $\text{MnFe}\mathcal{X}\mathcal{Y}$ was done in the M direction. $\text{FeCo}\mathcal{X}\mathcal{Y}$ had the highest T_C and large M . $\text{MnFe}\mathcal{X}\mathcal{Y}$ had the highest M and moderate T_C . $\text{FeCoNi}\mathcal{X}$ had the largest T_C . $\text{MnFeCo}\mathcal{X}$ and $\text{MnFeNi}\mathcal{X}$ had the largest M . $\text{MnFe}\mathcal{X}\mathcal{Y}$ and $\text{MnFeNi}\mathcal{X}$ had the largest M . Notably, $\{\text{MnFeCo}\mathcal{X}\} \subset \{\text{MnFe}\mathcal{X}\mathcal{Y}\}$ or $\{\text{FeCo}\mathcal{X}\mathcal{Y}\}$.

$\text{FeCo}\mathcal{X}\mathcal{Y}$, $\text{MnFe}\mathcal{X}\mathcal{Y}$, $\text{MnFeCo}\mathcal{X}$, $\text{FeCoNi}\mathcal{X}$, and $\text{MnFeNi}\mathcal{X}$ are separately shown in Figs. 1 and 2 of Ref. [37]. There are materials with small M and T_C in both the BCC and FCC phases. As will be shown later, the reason for the small M is that the elements have an antiferromagnetic spin configuration.

Next, materials with the maximum values are listed. The top five systems with the highest M s were MnFeCoNi , MnFeCoRh , MnFeCoPt , MnFeCoIr , and MnFeCoAu in P_A or P_M in the BCC phase. Their values were 1.94, 1.93, 1.89, 1.88, and 1.76 T, respectively. The top five systems with the highest M s were MnFeNiAu , MnFeNiCu , MnFeCoTi , MnFeCoSn , and CrMnCoTi in P_A or P_M in the FCC phase. Their values were 1.52, 1.51, 1.49, 1.46, and 1.46 T, respectively. The

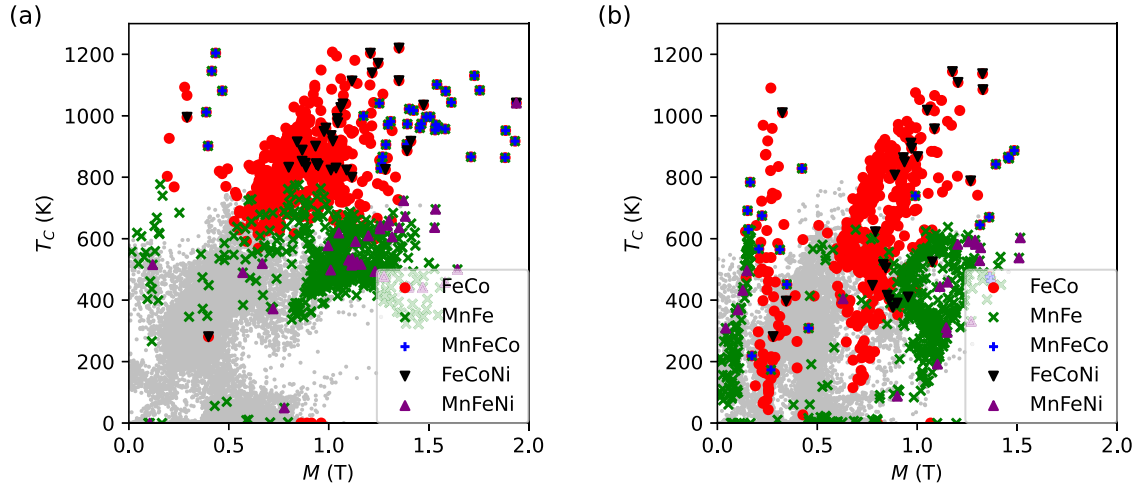


FIG. 4. M and T_C in P_M (a set of all magnetic materials) in the BCC phase (a) and in the FCC phase (b). The materials containing FeCo, MnFe, MnFeCo, FeCoNi and MnFeNi are labeled. P_M is defined in Eq. (5)

correlation between M and units of the spin moment per primitive cell is shown in Fig. 3 in Ref. [37] for reference. The top five systems with the highest T_C s were FeCoNiPd, FeCoCuAg, MnFeCoIn, FeCoNiAg, and FeCoCuAu in P_A or P_M in the BCC phase. Their values were 1220, 1207, 1204, 1204, and 1194 K, respectively. The top five systems with the highest T_C s were FeCoNiAg, FeCoNiPd, FeCoNiAu, FeCoCuHg, and FeCoNiPt in P_A or P_M in the FCC phase. Their values were 1144, 1137, 1109, 1090, and 1085 K, respectively. As shown in Figs. 3 and 4, the systems with BCC phase tend to have larger M and T_C , compared to those with FCC phase. This tendency can be also found in previous studies for high entropy alloys [38–40].

E. Residual electrical resistivity

The top five systems with the highest R_s were ScVCrTi, ScMnInRe, ScVYIn, ScVHfTi, and ScVYSn in P_A in the BCC phase. Their values were 213, 205, 204, 202, and 202 $\mu\Omega$ cm, respectively. The top five systems with the highest R_s were ScCrYBi, ScTiCrBi, ScTiCrPb, ScTiCrHg, and ScTiCrTi in P_A in the FCC phase. Their values were 186, 186, 184, 184, and 184 $\mu\Omega$ cm, respectively. Moreover, FeCoNiAg, FeCoCuAg, FeCoNiCu, CoNiCuAg, and FeNiCuAg had the lowest R in P_A in the BCC phase. Their values were 1, 2, 3, 5, and 5 $\mu\Omega$ cm, respectively. FeCoNiAg, FeCoCuAg, CoNiCuAg, FeNiCuAg, and FeCoNiPd had the lowest R in P_A in the FCC phase, with values as 1, 3, 5, 5, and 5 $\mu\Omega$ cm, respectively.

VI. DISCUSSION

A. Validation of lattice constants

In this calculation, the lattice parameter was the value calculated from the average value of the volume of the representative material in the experiment for each constituent element. Fig. 5 shows the labels of the BCC and FCC phases by Lederer *et al.* [41] for the total energy difference between FCC and BCC from first-principles calculations. The phase boundary threshold energy determined by decision tree regression to maximize the degree of separation of the distri-

bution by mean squared error (MSE) was 0.01995 Ry, which is indicated by the red vertical dotted line in the figure, where the total energy is the value per primitive unit cell. Of the 480 quaternary substances labeled BCC and FCC by Lederer *et al.* [41], the number of substances that were successfully calculated was 479, and the number of substances that failed to be classified at the energy threshold was 6. The classification accuracy was 0.950; thus, the classification model is valid. Although it is a side evidence, we were able to show the validation of the lattice constants. In the case of CrFeCoNi with FCC phase, the experimentally observed lattice constant is 3.5665 Å [42], while the lattice constant estimated from the atomic volume of each constituent element is 3.5789 Å.

In the experiment, when the solid solution phase appeared, the intermetallic phase might be observed simultaneously. Since the simultaneous presence of solid solution and inter-

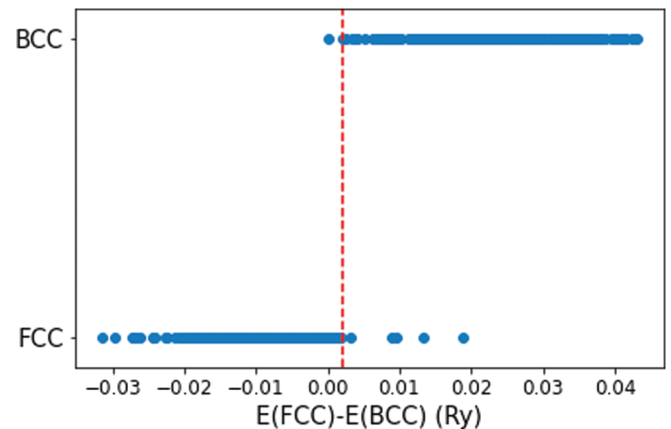


FIG. 5. Binary classification of the stable structures in 480 quaternary high entropy alloys. The vertical axis indicates the label of the experimentally confirmed structure (FCC or BCC) by Lederer *et al.* [41] The horizontal axis is the total energy difference between the FCC and BCC phases by the AkaiKKR code. The threshold value of 0.01995 Ry is determined by decision tree regression and is indicated by the vertical dashed line.

metallic phases is advantageous for mechanical strength, one strategy is to search for materials in which both phases tended to appear [43,44]. For example, for FeCoMnNi, the energy difference [$E(\text{BCC}) - E(\text{FCC})$] in this study was -0.0097 Ry, which was on the side of 0.02965 Ry FCC when the phase boundary threshold energy was added. However, Ha *et al.* [45] recently reported that FeCoMnNi had the BCC phase in thin films, implying that other materials might yield phases different from those predicted by *ab initio* calculations depending on the experimental setting, although 0.02965 Ry per primitive cell was a large value. Therefore, in this study, we did not consider the stability of the BCC and FCC phases but performed exhaustive calculations for each phase separately.

B. Characteristic features in (M, T_C) space

For areas where the distribution was sparse, their characteristic features could be easily identified using a scatter plot, but not for areas where the distribution was dense. Here, we identify the characteristic features of the parcels in 2D physical property space, (M, T_C) in P_M . It may be a good choice to execute density-based spatial clustering of applications with noise (DBSCAN) [46,47] for density clustering. However, our purpose is not clustering, but to identify the characteristic features of the dataset in areas. Our dataset was large, thus frequent itemset mining was useful and efficient because it could extract common elements quickly. In this section, an area was defined by digitizing its physical properties for simplicity. We called the area parcel and then used frequent itemset mining to identify characteristic features of the parcel.

We explain frequent itemset mining briefly. Frequent itemset mining uses the transaction database, which is made of a transaction. A transaction is made of items. There are variants of frequent itemsets, e.g., closed and maximum itemsets. A closed frequent itemset mining was used in this study. It is explained according to Uno *et al.* [48] as follows. For itemset P , $T(P)$ is the set of the transactions having P . The frequency of P is the number of transactions having P . For any two itemsets P and Q , $T(P \cup Q) = T(P) \cap T(Q)$ holds, and if $P \subseteq Q$ then $T(Q) \subseteq T(P)$. An itemset P is called closed if no other itemset Q satisfies $T(P) = T(Q)$, $P \subseteq Q$, i.e., if it is included in no other itemset of the same frequency. In other words, a closed itemset has the advantage of outputting items as the maximum set of items. There are many implementations of frequent itemset mining. In this analysis, LCM version 5.3 was used [49–52].

A transaction corresponded to data of material in this case. It comprised three components: (A) constituent elements of material: element1-element4; (B) a parcel ID by digitizing (M, T_C) of each material (Fig. 6); and (C) spin configuration of each material.

For (C), first, the local magnetizations m_x were sorted in descending order as $|m_1| > |m_2| > |m_3| > |m_4|$. We redefined m_1 as positive, and defined $m_i \geq m_{\text{th},1}$ as having a magnetic spin moment, and the ferromagnetic and antiferromagnetic spin configurations hierarchically in terms of m_i . For example, if $|m_1| > |m_2| > |m_3| > m_{\text{th},1} > |m_4|$ and $m_2 < 0$, $m_3 > 0$, the spin configuration items of this material were hierarchically defined as the strings FA, FAF, FAFN. For $|m_1| > |m_2| > m_{\text{th},1} > |m_3| > |m_4|$ and $m_2 < 0$, only FA and FAN

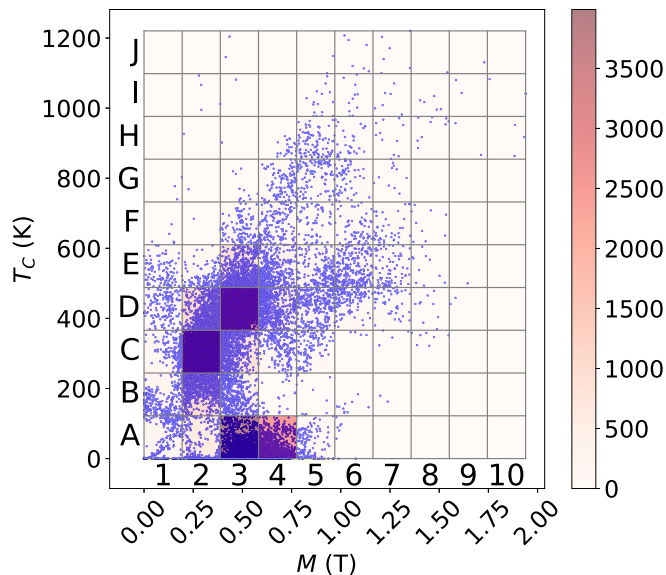


FIG. 6. Heatmap shows the frequency of each of the 10 divisions of M and T_C along with the scatter plot in the BCC phase.

were used. Notably, FANN was not adopted since its existence was obvious because $m_{\text{th},1} > |m_3| > |m_4|$. As a result, for example, CrRuHgBi had a transaction of Cr, Ru, Hg, Bi, parcel D4, FA, and FAN, where the last three strings were spin configurations.

For each parcel, itemsets with a frequency of at least 10 and at least 50% of the maximum frequency of each parcel were employed for analysis. Table II shows an example of a closed frequent itemset when the spin configuration was set to $m_{\text{th},1} = 0.2$. (a) is the result of parcel E5 in the BCC phase. The FF existed as many as the total number of transactions, 172: all materials had the FF spin configuration. The frequency of both Mn and FF was 145: the number of materials containing Mn with the FF spin configuration was 145. Similarly, the number of materials containing Mn and Co with the FF spin configuration was 117. The number of materials with the FFA and FF spin configuration was 89. The number of materials containing Mn with the FFA and FF spin

TABLE II. (a) Closed frequent itemset of the parcel (E5) in the BCC phase. The total number of transactions in the parcel is 172. (b) The closed frequent itemset of the parcel (E8) in the BCC phase. The total number of transactions in the parcel is 10.

(a)	Items	Frequency
1	FF	172
2	Mn, FF	145
3	Co, Mn, FF	117
4	FFA, FF	89
5	Mn, FFA, FF	87
(b)	Items	Frequency
1	Mn, FF, FFF	10
2	Fe, Mn, FFF, FF	8
3	Fe, Mn, FFFN, FFF, FF	7
4	Rh, Mn, FFF, FF	5

configuration was 87. Because $FFA \subset FF$, which is a physical interpretation and is beyond the itemset mining operation, Mn, FFA and FF were interpreted as materials containing Mn with the FFA spin configuration.

(B) is the closed frequent itemset of parcel E8. The materials containing Mn and FFF were 10. The materials containing Fe, Mn, and FFF were eight. The materials containing Fe, Mn, and FFFN are seven. The materials containing Rh, Mn, and FFF are five. For reference, the transactions in parcel E8 are shown in Table I of Ref. [37]. In Fig. 7 (see below), for example, when both Co and Mn were included, they were connected by an underscore: Co_Mn.

First, the number of transactions in each parcel, which was digitized by dividing M and T_C into 10 parts, respectively, are shown in Fig. 7(a) in the BCC phase. The parcels were colored according to their number of transactions and shaded when the number was less than 10. Parcels A3, A4, C2, and D3 had particularly large numbers of materials—3991, 2452, 3230, and 2930 transactions, respectively. In the following, we examine the constituent elements and spin configurations.

Figure 7(b) shows the combination of elements in each plot with letters. The elemental features of parcel E5 corresponding to Table II(a) were Mn and Co_Mn. The elemental features were Mn, Fe_Mn, and Rh_Mn in parcel E8 corresponding to Table II(b). Figure 7(c) shows the spin configuration with the highest frequency in each parcel. For example, parcel E5 was FF, and cell E8 was FFF. Notably, the magnetization was sorted in the descending order of absolute value, which did not correspond to the order of the elements shown in Fig. 7(b) because there were third and fourth elements.

As shown in Fig. 7(b), Fe, Co, and Mn were the most frequently detected elements. The lower right region of the figure, where M was large and T_C was small, was dominated by Mn, whereas, the upper left region, where M was small and T_C was large, was dominated by Fe and Co. Parcel D3 was Fe-dominated, but the spin configuration above 50% was unidentified, which meant that a few spin configurations were mixed. Parcel A3 was Mn-dominated, and the spin configuration was FA. Parcel A4 had the same Mn dominance, but there was no majority in spin configurations, or a few spin configurations were mixed.

This is a repeat of the previous example; we now consider parcels where the number of transactions is small but M or T_C are large. For example, parcel I6, which had a high T_C , had the FF spin configuration and comprised mainly $CoFe\mathcal{X}\mathcal{Y}$, a material containing both Co and Fe. Parcel E8, which had a large M and an FFF spin configuration comprised mainly $Mn\mathcal{X}\mathcal{Y}Z$, $FeMn\mathcal{X}\mathcal{Y}$, and $MnRh\mathcal{X}\mathcal{Y}$. Refer to Table I of Ref. [37] for the corresponding items.

We consider changes in neighboring Parcel. For example, from parcel A3 to A5, there was a change from a predominantly Mn and FA configuration in A3 to a predominantly Mn and mixed spin configurations in A4, and then to a predominantly Mn and Ni and FF spin configuration in A5. Next, consider parcels E2–E4, Fe had the same predominance, but different spin configurations were observed: parcel E2 was FA, parcel E3 was interspersed with a few spin configurations, and parcel E4 was FF. This is a change in the spin configuration of the second through fourth elements. In ad-

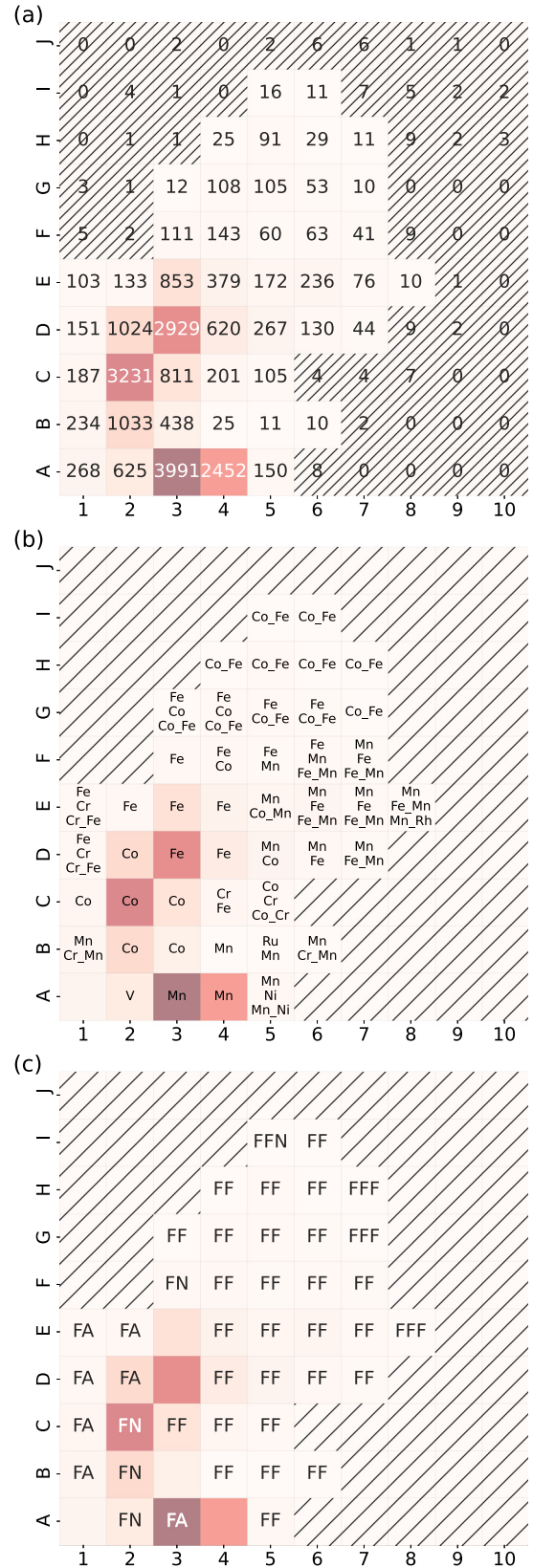


FIG. 7. Main features of parcels where M and T_C are divided into 10 parts in the BCC phase. (a) shows the number of transactions, (b) shows the elemental features, and (c) shows the most frequent spin configuration. The parcels are colored corresponding to their numbers of transactions and shaded when they are less than 10.

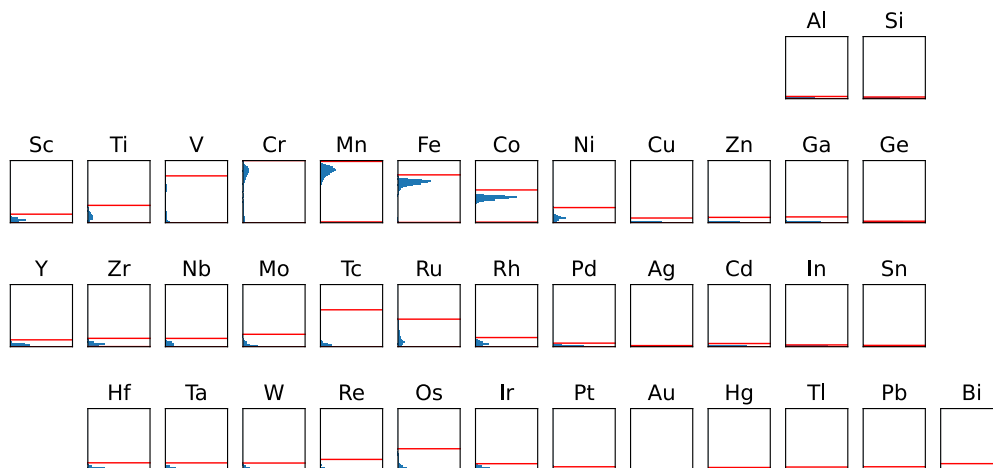


FIG. 8. Distribution of $|m(\mathcal{W})|$ in the BCC phase. The vertical axis represents $|m(\mathcal{W})|$ and the horizontal axis represents the occurrence. Each panel is arranged according to the position of the element on the periodic table. The detailed explanation is given in the main body.

dition, the neighboring F4 was dominated by not only Fe but also Co. The scatter plots of FA, FN, and FF spin configurations for Mn, Fe, and Co, respectively, are shown in Fig. 4 of Ref. [37], where the alpha transparency must be tuned to reproduce the above knowledge.

The same analysis for the FCC phase is shown in Figs. 5 and 6 of Ref. [37]. In the 10×10 parcels, the area where FN was dominant was larger in the FCC phase. The coarse-graining of 10 divisions in a 2D space of (M, T_C) was performed here to illustrate with letters, but it was possible to create fine parcels in a 3D space of $(M, T_C, \text{electrical resistivity})$. According to the illustration using a scatter plot, the impression may change depending on how it is drawn, but it is possible to identify the features quantitatively and quickly by frequent itemset mining.

C. Local magnetic moment

Magnetic properties can be discussed in terms of the spin magnitude of each element and the spin correlation between elements. First, the spin magnitude of each element is discussed, followed by the spin correlation between elements.

Figure 8 shows the distribution of $|m(\mathcal{W})|$ in the BCC phase. The vertical axis represents $|m(\mathcal{W})|$, and the horizontal axis represents the occurrence. Each panel was arranged according to the position of the element on the periodic table. The vertical and horizontal scales of each panel were the same. The scale of the vertical axis was from 0 to $4.492 \mu_B$. The red horizontal lines indicate the maximum and minimum values of each distribution. If the red horizontal lines coincide with the upper and lower frames, the horizontal lines of the maximum and minimum values are invisible.

Although V and Tc favored antiparallel spin configuration for Mn, Fe, Co, and Ni, Os often had parallel spin configuration, thus expected to have high total magnetization and a large contribution to T_C . Unfortunately, as will be shown later, the contribution to the maximum total magnetization of Os was not large.

The elements, \mathcal{X} , in E_M could have $|m(\mathcal{X})|$ close to zero depending on the material, which is shown as lower horizontal lines. However, the peaks of the maximum values were mainly

at large finite values: the magnetization was finite regardless of the constituent elements. This agreed with the knowledge of magnetic materials. Moreover, V, Tc, Ru, and Os showed relatively large changes depending on the element combination. The results were the same in the FCC phase (Fig. 8 in Ref. [37]).

D. Two-body spin-spin correlations

The Slater-Pauling curve [53,54] is famous for the total magnetization of alloys. The spin correlations between elements are known to be antiparallel in the local spins between early and late transition metals and parallel in the spins between late transition metals. We have already shown that some spin configurations are mixed in the (M, T_C) space. Here, we define the number of two-body parallel (F) and antiparallel (AF) spin configurations between elements as follows:

$$n_{\mathcal{X}\mathcal{Y}}(\text{AF}) = \sum_{\{\mathcal{X}, \mathcal{Y} | \mathcal{X}, \mathcal{Y} \in \mathcal{P}_2(E_\lambda)\}} \theta(|m_{\mathcal{X}}| > m_{\text{th},1}) \times \theta(|m_{\mathcal{Y}}| > m_{\text{th},1}) \theta(m_{\mathcal{X}} m_{\mathcal{Y}} < 0), \quad (7)$$

$$n_{\mathcal{X}\mathcal{Y}}(\text{F}) = \sum_{\{\mathcal{X}, \mathcal{Y} | \mathcal{X}, \mathcal{Y} \in \mathcal{P}_2(E_\lambda)\}} \theta(|m_{\mathcal{X}}| > m_{\text{th},1}) \times \theta(|m_{\mathcal{Y}}| > m_{\text{th},1}) \theta(m_{\mathcal{X}} m_{\mathcal{Y}} > 0), \quad (8)$$

where $\theta(x)$ is the Heaviside step function.

The $n_{\mathcal{X}\mathcal{Y}}(\text{AF})$ and $n_{\mathcal{X}\mathcal{Y}}(\text{F})$ of the BCC phase are shown in Fig. 9. (a) denotes $n_{\mathcal{X}\mathcal{Y}}(\text{AF})$; Ti, V, Mn, and Fe in $3d$, Zr, Nb, Mo, Tc, and Ru in $4d$, and Ta, W, Re, and Os in $5d$ have large numbers of antiparallel spin configurations for Cr; Sc, Ti, V, and Cr in $3d$, Y, Zr, Nb, Mo, and Tc in $4d$, and Hf, Ta, W, and Re in $5d$ had large numbers of antiparallel spin configurations for Mn. Fe had a large number of antiparallel spin configurations for Sc, Ti, V, and Cr in $3d$, Y, Zr, Nb, and Mo in $4d$, and Hf, Ta, and W in $5d$. Co had a relatively large number of Sc, Ti, and V in $3d$, Y and Zr in $4d$, and Hf in $5d$, oriented in antiparallel spins. As the elements changed from Cr to Mn to Fe to Co, the region of elements showing antiparallel spin configurations narrowed to the upper left side of the periodic table.

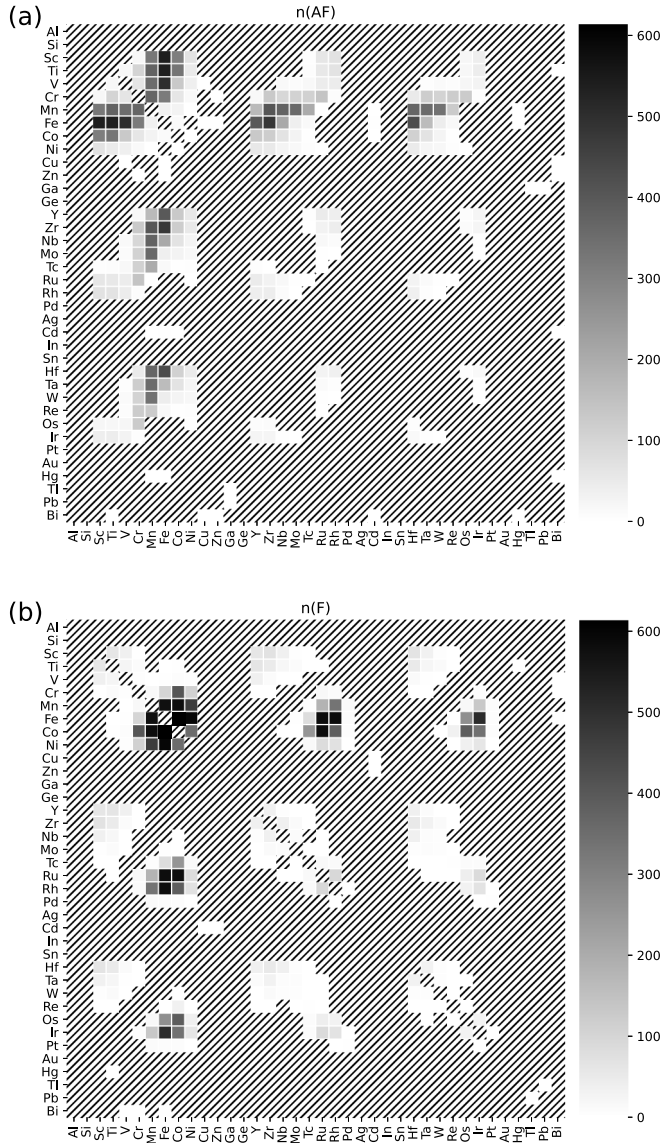


FIG. 9. (a) $n_{\mathcal{X}\mathcal{Y}}(\text{AF})$ and (b) $n_{\mathcal{X}\mathcal{Y}}(\text{F})$ when $m_{\text{th},1} = 0.2 \mu_B$ in the BCC phase. See also the definition in the main text.

(b) shows $n_{\mathcal{X}\mathcal{Y}}(\text{F})$; for Cr, there were large numbers of parallel spin configurations of Fe, Co, and Ni; for Mn, in addition to Fe, Co, and Ni, there were large numbers of parallel spin configurations of Ru, Rh, and Ir in $4d$ and $5d$; for Fe, there were large numbers of parallel spin configurations of Mn, Co, and Ni in $3d$, Rh and Ph in $4d$, and Os and Ir in $5d$; for Co, there were large numbers of parallel spin configurations of Cr, Mn, Fe, and Ni in $3d$, Tc, Rh, and Ru in $4d$, and Os and Ir in $5d$. With the change from Cr to Mn, Fe, and Co, the area of elements showing parallel spin configurations widened to the left and right of the group and the bottom of $4d$ and $5d$ in the periodic table.

The same analysis for FCC is shown in Fig. 9 of Ref. [37]. Though FCC had slightly less magnetic material, there was a similar trend for the FCC and BCC phases.

E. Effective additive elements to the magnetic properties

1. Definitions

Magnetic binary alloys and magnetic impurities have been studied [53–57]. It is well known that alloy systems, including $3d$ transition metals with both localized and itinerant characteristics, such as Cr, Mn, Fe, Co, and Ni, exhibit interesting magnetic properties. Consider the situation where magnetic impurities are doped into ferromagnetic Fe, Co, and Ni. When the dopants are early transition metals, such as Sc, Ti, V, and Cr, the impurity induced narrow virtual bound states were generated around E_F due to the repulsing potentials. These virtual bound states were negatively polarized in the opposite direction to the parent element, thereby decreasing the total magnetic moment. Mn is a critical atom in which the direction of the magnetic moment is sensitive to the parent lattice constant and exchange-correlation potential [55–57].

Further, Pd is known to increase the local magnetization of nearby Ni atoms [58] and is expected to improve magnetic properties as an additive element. In fact, a material search including Rh, which is a neighboring element of Pd, has been conducted [59]. In this section, we analyze how much E_N improves the physical properties of E_M .

Closed frequent itemset mining of (M, T_C) identified high-frequency regions. In this section, we identify the low-frequency regions. We define $E_W \in E_N$, and the contribution of E_W to the elements $\{\mathcal{X}, \mathcal{Y} \mid \mathcal{X}, \mathcal{Y} \in \mathcal{P}_2(E_M)\}$ is evaluated by

$$\Delta F_2(\mathcal{X}, \mathcal{Y}; E_W) = F_2(\mathcal{X}, \mathcal{Y}) - F_2^{\text{loo}}(\mathcal{X}, \mathcal{Y}; E_W), \quad (9)$$

where

$$F_2(\mathcal{X}, \mathcal{Y}) = \max_{\mathcal{Z}, \mathcal{W} \in \mathcal{P}_2(E_N)} F(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W}), \quad (10)$$

$$F_2^{\text{loo}}(\mathcal{X}, \mathcal{Y}; E_W) = \max_{\mathcal{Z}, \mathcal{W} \in \mathcal{P}_2(E_N - E_W)} F(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W}), \quad (11)$$

and superscript *loo* is an abbreviation for leave-one-out.

Function $F(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W})$ was assumed to be M or T_C of the material, $\mathcal{X}\mathcal{Y}\mathcal{Z}\mathcal{W}$. Similarly, the equations were defined when three and one elements were chosen from E_M and E_N , respectively, without duplication and when one and three elements were chosen from E_M and E_N , respectively, without duplication as

$$\Delta F_3(\mathcal{X}, \mathcal{Y}, \mathcal{Z}; E_W) = F_3(\mathcal{X}, \mathcal{Y}, \mathcal{Z}) - F_3^{\text{loo}}(\mathcal{X}, \mathcal{Y}, \mathcal{Z}; E_W), \quad (12)$$

$$\Delta F_1(\mathcal{X}; E_W) = F_1(\mathcal{X}) - F_1^{\text{loo}}(\mathcal{X}; E_W). \quad (13)$$

2. Magnetization

Figure 10(a) shows $M_1(\mathcal{X})$, $M_2(\mathcal{X}, \mathcal{Y})$, and $M_3(\mathcal{X}, \mathcal{Y}, \mathcal{Z})$ in the BCC phase. The horizontal axis represents the combination from $\mathcal{P}_k(E_M)$. When $\{\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W} \mid \mathcal{X} \in \mathcal{P}_1(E_M), \mathcal{Y}, \mathcal{Z}, \mathcal{W} \in \mathcal{P}_3(E_N)\}$, $\{\text{Cr}\}$, $\{\text{Fe}\}$, and $\{\text{Mn}\}$; when $\{\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W} \mid \mathcal{X}, \mathcal{Y} \in \mathcal{P}_2(E_M), \mathcal{Z}, \mathcal{W} \in \mathcal{P}_2(E_N)\}$, $\{\text{Mn, Fe}\}$, $\{\text{Mn, Co}\}$, $\{\text{Fe, Co}\}$, and $\{\text{Cr, Mn}\}$; and when $\{\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W} \mid \mathcal{X}, \mathcal{Y}, \mathcal{Z} \in \mathcal{P}_3(E_M), \mathcal{W} \in \mathcal{P}_1(E_N)\}$, $\{\text{Mn, Fe, Co}\}$, $\{\text{Mn, Fe, Ni}\}$, and $\{\text{Cr, Mn, Fe}\}$ exhibit relatively large total magnetization.

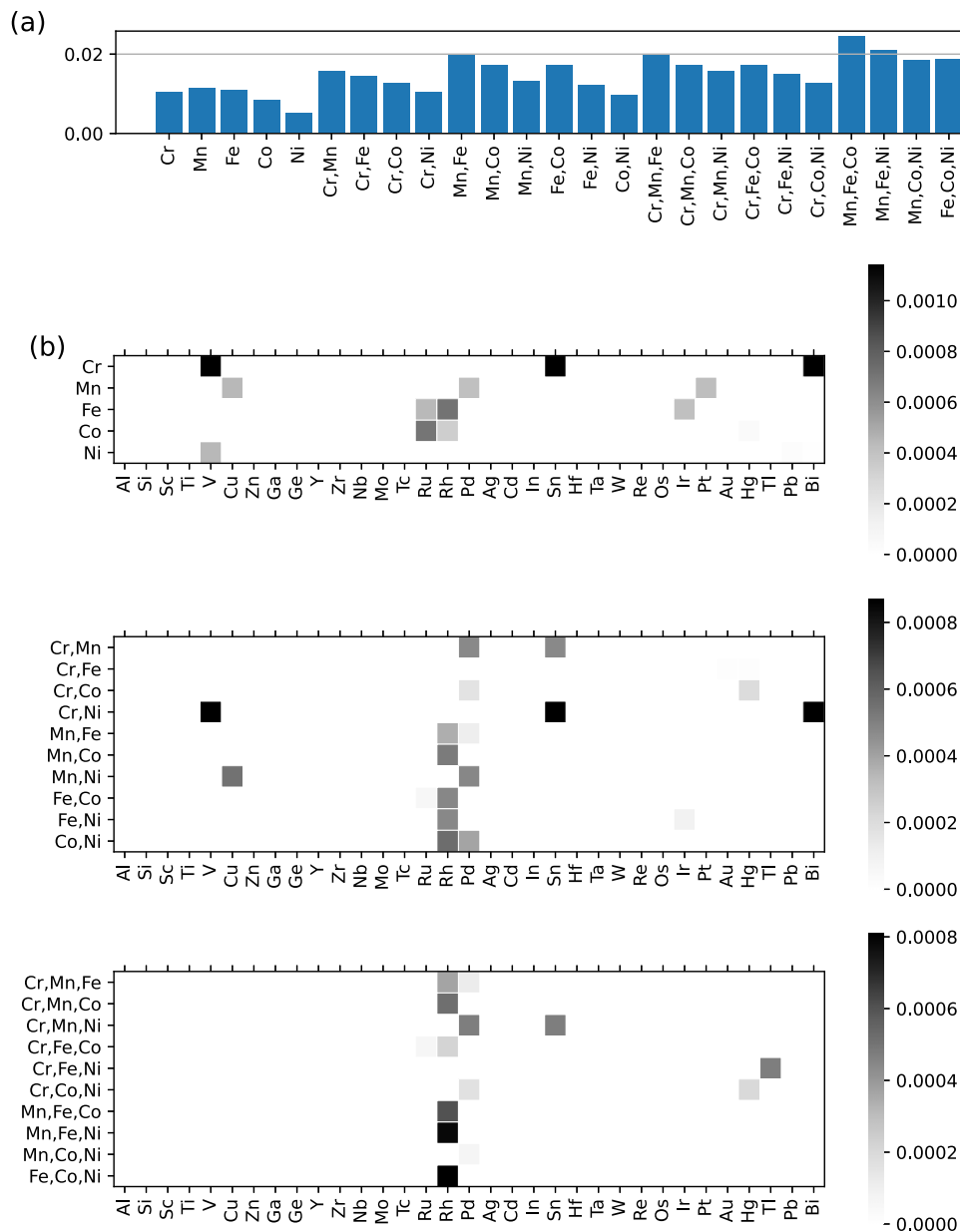


FIG. 10. (a) Combination from $\mathcal{P}_k(E_M)$ (in the horizontal axis) versus $M_k(E_M)$ (in the vertical axis) are plotted in the BCC phase. The unit of the vertical axis is μ_B/Bohr^3 . (b) Heatmaps of $\Delta M_1(\mathcal{X}; E_W)$, $\Delta M_2(\mathcal{X}, \mathcal{Y}; E_W)$, and $\Delta M_3(\mathcal{X}, \mathcal{Y}, \mathcal{Z}; E_W)$ (in the vertical axis) in the BCC phase, where $E_W := \{\mathcal{W} \mid \mathcal{W} \in E_N\}$ (in the horizontal axis). $M_k(\mathcal{X})$ and $\Delta M_k(\mathcal{X})$ are defined in Eqs. (9)–(13) in Sec. VI E 1. The examples of the combinations from $\mathcal{P}_k(E_M)$ are shown in Sec. VI E 2.

Figure 10(b) shows $\Delta M_1(\mathcal{X}; E_W)$, $\Delta M_2(\mathcal{X}, \mathcal{Y}; E_W)$, and $\Delta M_3(\mathcal{X}, \mathcal{Y}, \mathcal{Z}; E_W)$, where $E_W := \{\mathcal{W} \mid \mathcal{W} \in E_N\}$. The vertical axis represents $\mathcal{P}_k(E_M)$. The horizontal axis represents E_W . We will examine the case of large total magnetization in the previous paragraph. The largest contributions to $\Delta M_1(\text{Cr}; E_W)$ is given by $E_W = \{\text{V}\}$, $\{\text{Sn}\}$, and $\{\text{Bi}\}$. The largest contribution to $\Delta M_2(\text{Mn, Fe}; E_W)$, $\Delta M_2(\text{Mn, Co}; E_W)$, and $\Delta M_2(\text{Fe, Co}; E_W)$ is given by $E_W = \{\text{Rh}\}$. The largest contribution to $\Delta M_2(\text{Cr, Mn}; E_W)$ is given by $E_W = \{\text{Pd}\}$ and $\{\text{Sn}\}$. The largest contribution to $\Delta M_3(\text{Mn, Fe, Co}; E_W)$, $\Delta M_3(\text{Mn, Fe, Ni}; E_W)$, and $\Delta M_3(\text{Cr, Mn, Fe}; E_W)$ is given by $E_W = \{\text{Rh}\}$.

Rh and Pd contribute significantly, but, as mentioned above, V, Sn, Bi, and Ti also contribute significantly. Notably,

the analysis of Fig. 10 is for the case where only one element was selected from E_W , and for example, $\Delta M_1(\mathcal{X}; E_W)$ might be lowered when two elements were selected from E_W . Notably, the analysis was vulnerable when there were anomalous large values at the highest physical quantity. The same analysis for the FCC phase is performed in Fig. 10 of Ref. [37].

3. Curie temperature

Takahashi *et al.* proposed, by first-principles calculations, that T_C of binary alloys possesses somewhat similar behavior to the Slater-Pauling curve [25]. They showed that introducing a small amount of early transition metals, such as V and Cr, to

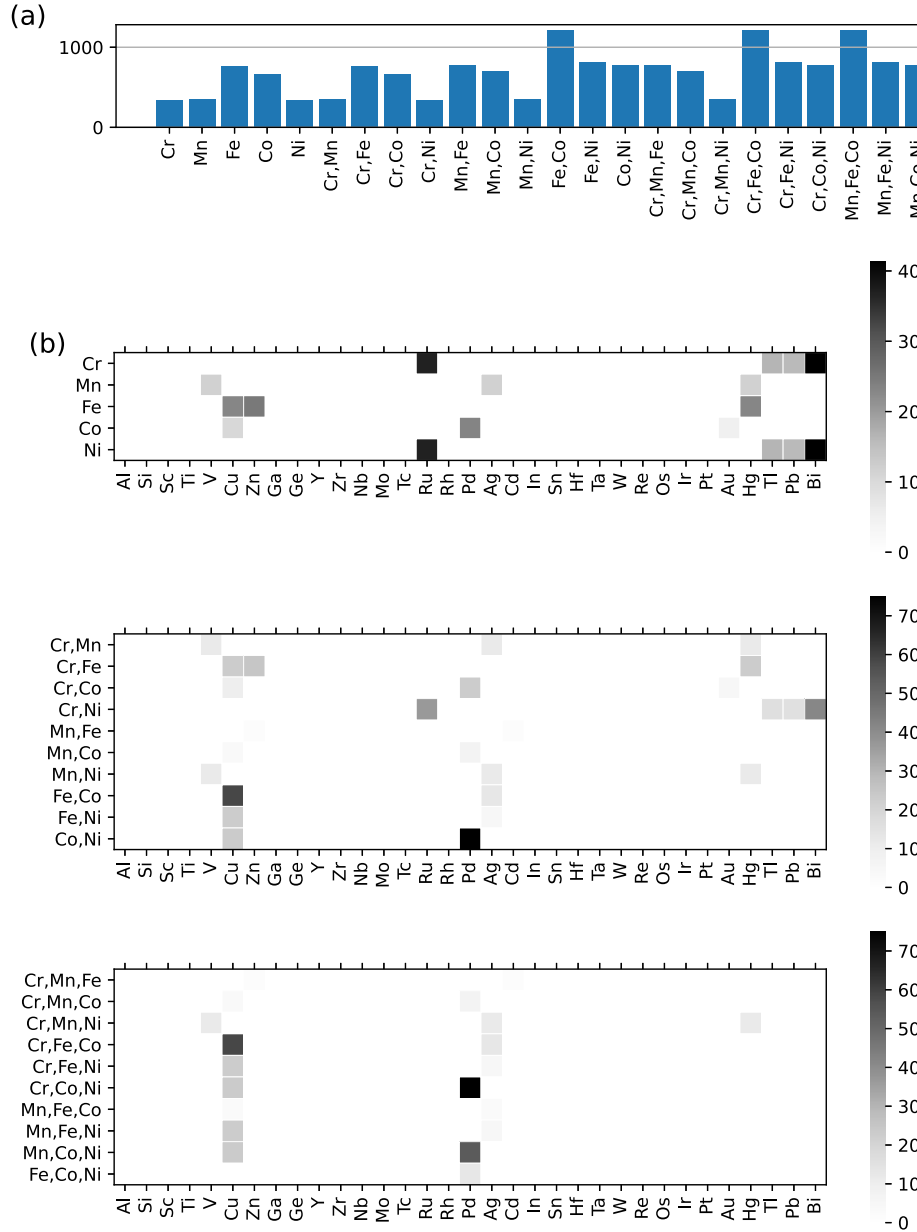


FIG. 11. (a) Vertical axis shows $T_{C,1}(\mathcal{X})$, $T_{C,2}(\mathcal{X}, \mathcal{Y})$, and $T_{C,3}(\mathcal{X}, \mathcal{Y}, \mathcal{Z})$. The horizontal axis is the combination from $\mathcal{P}_k(E_M)$. The unit of the vertical axis is K. (b) Heat maps of $\Delta T_{C,1}(\mathcal{X}; E_W)$, $\Delta T_{C,2}(\mathcal{X}, \mathcal{Y}; E_W)$, and $\Delta T_{C,3}(\mathcal{X}, \mathcal{Y}, \mathcal{Z}; E_W)$ (in the vertical axis) in the BCC phase, where $E_W := \{\mathcal{W} \mid \mathcal{W} \in E_N\}$ (in the horizontal axis). $T_{C,k}(\mathcal{X})$ and $\Delta T_{C,k}(\mathcal{X})$ are defined in Eqs. (9)–(13) in Sec. VI E 1. The examples of the combinations from $\mathcal{P}_k(E_M)$ are shown in Sec. VI E 2.

BCC-Fe increased T_C . Here, the effect to T_C was examined in HEA4, the atomic concentration of which was as much as (i.e., 25%), larger than of Ref. [25], where the maximum doping was 10%.

Figure 11(a) shows $T_{C,1}(\mathcal{X})$, $T_{C,2}(\mathcal{X}, \mathcal{Y})$, and $T_{C,3}(\mathcal{X}, \mathcal{Y}, \mathcal{Z})$ in the vertical axis. The horizontal axis represents $\mathcal{P}_k(E_M)$. $\{\text{Fe,Co}\}$ for $\{\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W} \mid \mathcal{X}, \mathcal{Y} \in \mathcal{P}_2(E_M), \mathcal{Z}, \mathcal{W} \in \mathcal{P}_2(E_N)\}$, and $\{\text{Cr,Fe,Co}\}$, $\{\text{Mn,Fe,Co}\}$, and $\{\text{Fe,Co,Ni}\}$ for $\{\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W} \mid \mathcal{X}, \mathcal{Y}, \mathcal{Z} \in \mathcal{P}_3(E_M), \mathcal{W} \in \mathcal{P}_1(E_N)\}$ show large T_C s.

Figure 11(b) shows $\Delta T_{C,1}(\mathcal{X}; E_W)$, $\Delta T_{C,2}(\mathcal{X}, \mathcal{Y}; E_W)$, and $\Delta T_{C,3}(\mathcal{X}, \mathcal{Y}, \mathcal{Z}; E_W)$. $\mathcal{P}_k(E_M)$ is shown in the vertical axis and E_W in the horizontal axis. The analysis was performed for materials that showed large T_C s in the previous paragraph.

The largest contributions to $\Delta T_{C,1}(\text{Fe}; E_W)$ are given by $E_W = \{\text{Cu}\}$, $\{\text{Zn}\}$, and $\{\text{Hg}\}$. The largest contribution to $\Delta T_{C,1}(\text{Co}; E_W)$ is given by $E_W = \{\text{Pd}\}$. The largest contribution to $\Delta T_{C,2}(\text{Fe, Co}; E_W)$ is given by $E_W = \{\text{Cu}\}$. The largest contribution to $\Delta T_{C,3}(\text{Cr, Fe, Co}; E_W)$ is given by $E_W = \{\text{Cu}\}$. The largest contribution to $\Delta T_{C,3}(\text{Mn, Fe, Co}; E_W)$ is given by $E_W = \{\text{Ag}\}$. The largest contribution to $\Delta T_{C,3}(\text{Fe, Co, Ni}; E_W)$ is given by $E_W = \{\text{Pd}\}$. In summary, not only Pd but also Cu, Zn, Ag, and Hg significantly contributed to T_C s. The same analysis for the FCC phase is performed in Fig. 11 of Ref. [37].

It is well known that the mean-field approximation tends to overestimate T_C . There are several other approaches for estimating T_C in the scale-bridging method, such as the

cluster-like approximation [21,22], Tyablicov approximation [23,24], and Monte Carlo simulation. In particular, the Tyablicov approximation and Monte Carlo simulation, which can consider a magnetic percolation effect, can give a rather accurate estimation. Reference [4] considered that the present chemical trend of T_C changes depending on the evaluating methods. It is worthwhile to compare our exhaustive calculation and previous studies. The experimentally observed T_C of CrFeCoNi with FCC phase is 130 K [42]. There also exists the previous study for the estimation of T_C by exact muffin-tin orbital (EMTO)-CPA method, where the obtained T_C is 156 K [60]. On the other hand, our calculated value of T_C is 423 K. One of the reasons why our calculated value is larger than the experimental result might be an error in the mean field approximation. Also a possible reason is that J_{ij} in itinerant systems generally depends on temperature. As for this point, in the previous EMTO-CPA study the T_C was estimated by the mean-field approximation based on the total energy difference between ferromagnetic and local moment disordered (LMD) states. In our exhaustive calculation, T_C is estimated by combining the mean-field approximation with J_{ij} calculated in a ferromagnetic CPA medium. The discrepancy between our result and the EMTO-CPA-LMD study comes from the effects of the spin scatterings. In the present material space, there are systems which do not have LMD solutions (e.g., systems with small local magnetic moments). Therefore we performed the exhaustive calculations for T_C using J_{ij} calculated in a ferromagnetic CPA medium.

F. Residual electrical resistivity

This study has exhaustive data, and data-driven science is useful for testing hypotheses. The electrical resistivity of magnetic materials is generally explained as a combination of a large d -electron partial density of states (PDOS) and spin disorder, i.e., a small d -electron PDOS with a majority spin and a large d -electron PDOS with a minority spin, which results in a large (residual) electrical resistivity [61–63]. However, it was impossible to give a valid prediction model using PDOS as an explanatory variable for all materials, which might be similar to the situation where, for example, the superconducting transition dependence on the doping level was discussed as intrinsic in Cuprate high-temperature superconductors scales among different materials [64]. However, when looking for correlations for the superconducting transition temperature among all materials, changes in materials, such as LSCO and YBCO, are more important for the superconducting transition temperature.

Further, the regression model of a material series that changes only \mathcal{W} in \mathcal{XYZW} had some predictive performance values, but the PDOS of the d electrons with large minority spin was not an important explanatory variable. The reason PDOS, which takes a k sum, was not a good explanatory variable was that the height of the peak or imaginary part corresponding to the eigenenergy of the Green function of CPA on the Fermi level greatly differed for different k points. It may be understood that the k point that significantly contributes to electrical conduction is only a part of the Fermi level [18,65]. Notably, there were cases in which the large PDOS of d electrons and spin disorder can be qualitatively

TABLE III. Regression methods for the resistivity, method of making explanatory variables, and R_{test}^2 in the BCC phase are shown. X1 means the explanatory variables as they are, and X2 and X3 mean the squared and cubed explanatory variables are added to the explanatory variables.

Regression method	Explanatory variables	R_{test}^2
Linear regression	X1	0.754
Linear regression	X1, X2	0.815
Linear regression	X1, X2, X3	0.840
Random forest regression	X1	0.964
k-neighbors regression	X1	0.941

explained as contributing to the large electrical resistance. The results of the data analysis showed that numerous cases could not be regressed using the conventional explanations because the correlation was determined by a large number of characteristics of the data instances; most cases could not be regressed using the conventional explanation alone.

In the previous paragraph, the explanatory variable was PDOS, but in this paragraph, the regression is conducted using different explanatory variables. We used 40 explanatory variables made by the periodic table properties, (1) group, (2) row, and elemental physical properties, (3) electronegativity, (4) atomic radius, (5) atomic radius calculated, (6) thermal conductivity, (7) boiling point, (8) melting point, (9) molar volume, and (10) logarithm of electrical resistivity by applying (mean, standard deviation, min, max) to satisfy the elemental order invariance. Although these exhaustive calculation data were for isometric HEA, it was expected that these explanatory variables could also be used for HEA close to isometric HEA, and these explanatory variables could be reasonably applied to unknown data other than these data, so the prediction problem was considered. In the following, the explanatory variable, for example, generated by group and std, which is an abbreviation of the standard deviation are connected by underscores and denoted as `group_std`.

Table III shows the coefficient of determination (R_{test}^2) obtained by dividing R observations in P_A for the BCC phase into single training and test datasets at a ratio of 7:3, and using the test dataset to obtain a predictive evaluation index. Linear, random forest, and k -neighbor regressions were used as regression methods. In linear regression, up to the first, second, and third power of the explanatory variables were used. In linear regression, when we used up to the third power of the explanatory variables, we obtained about $R_{\text{test}}^2 = 0.840$. Moreover, for the random forest and k -neighbors regressions, we obtained $R_{\text{test}}^2 = 0.964$ and 0.941 , respectively. Although the high predictive performance of the random forest regression might be an artifact of the decomposition tree, we showed that the k -neighbor regression using correlation (similarity) had R_{test}^2 comparable to the random forest regression. The same analysis for the FCC phase is performed in Table II of Ref. [37], which gives a slightly lower prediction performance but a reasonable enough prediction model. The observed versus predicted R values of the test set by a k -neighbor regression is depicted in Figs. 13 and 14 of Ref. [37].

Here, the number of neighbors and explanatory variables of the k -neighbor regression are optimized as follows. First,

the only evaluation measure in supervised learning is the agreement of the supervised data with the target variables, and the importance of the explanatory variable was evaluated from the decrease in R^2_{test} by permutation importance, which randomized the explanatory variables. For the optimization of explanatory variables, we gradually excluded an explanatory variable with the lowest permutation importance and obtained explanatory variables and R^2_{test} in the BCC phase (Table III of Ref. [37]).

According to Table III in Ref. [37], the regression model with each number of explanatory variables contained group_* explanatory variables in many cases. Since this was no exhaustive model search, it might be an artifact of the selection algorithm for the explanatory variables. However, from these results, we can hypothesize that we could understand the trend of electrical conductivity from the periodic table variables. According to this hypothesis, we newly created explanatory variables by acting (mean, stdev, min, max) on the rows and columns of the periodic table, and conducted an exhaustive search of $2^8 - 1$ regression models with the k -neighbor regression, where the number of neighbors was optimized.

Regression models are approximate models, and one explanatory variable can often be supplemented by the combination of other explanatory variables. So, a regression model with another explanatory variable that is completely different from the best regression model may have almost the same evaluation index value. Therefore the explanatory variables important in a given regression model differ from those important in the set of (all) regression models. As methods to evaluate the important explanatory variables from the search model set, the following methods have been proposed. The method of arranging explanatory variables in the form of heat maps for evaluation index values [66–69], relevance analysis [70,71], which uses the leave-one-out and add-one-in approach, and subgroup relevance analysis [72,73], which defines explanatory variable groups using domain knowledge and performs the group leave-one-out, add-one-in approach, and evaluating the degree of separation of distributions [74]. Each method is suitable for different purposes. In this paper, strong relevance analysis and R^2_{test} distribution were used together, and closed frequent itemset mining was used to identify the common features of the area in the R^2_{test} distribution.

Figure 12(a) shows histograms of $R^2_A := \{\forall k, R^2_{\text{test}}(p_1, \dots, p_k) \mid p_1, \dots, p_k \in \mathcal{P}_k(V)\}$ and $R^2_W := \{\forall k, R^2_{\text{test}}(p_1, \dots, p_k) \mid p_1, \dots, p_k \in \mathcal{P}_k(V - V_W)\}$, where $V := \{\text{group, row}\} \times \{\text{min, max, ave, std}\}$ and $V_W := \{W \mid W \in V\}$. The evaluation index by strong relevance analysis [70,71] is defined as follows:

$$\Delta R^2(E_W) = \max_{\forall k, \{p_1, \dots, p_k\} \in \mathcal{P}_k(V)} R^2_{\text{test}}(p_1, \dots, p_k) - \max_{\forall k, \{p_1, \dots, p_k\} \in \mathcal{P}_k(V - V_W)} R^2_{\text{test}}(p_1, \dots, p_k). \quad (14)$$

When $V_W := \{\text{group_mean}\}$ or $\{\text{group_std}\}$, $\Delta R^2(E_W)$ was as large as 0.095 or 0.049, respectively.

To show the importance of group_mean and group_std, we analyzed the distribution histogram of R^2 again using a different method. In Fig. 12(a), several regions were separated by bin appear. The data in these regions could be interpreted as a transaction database where the explanatory

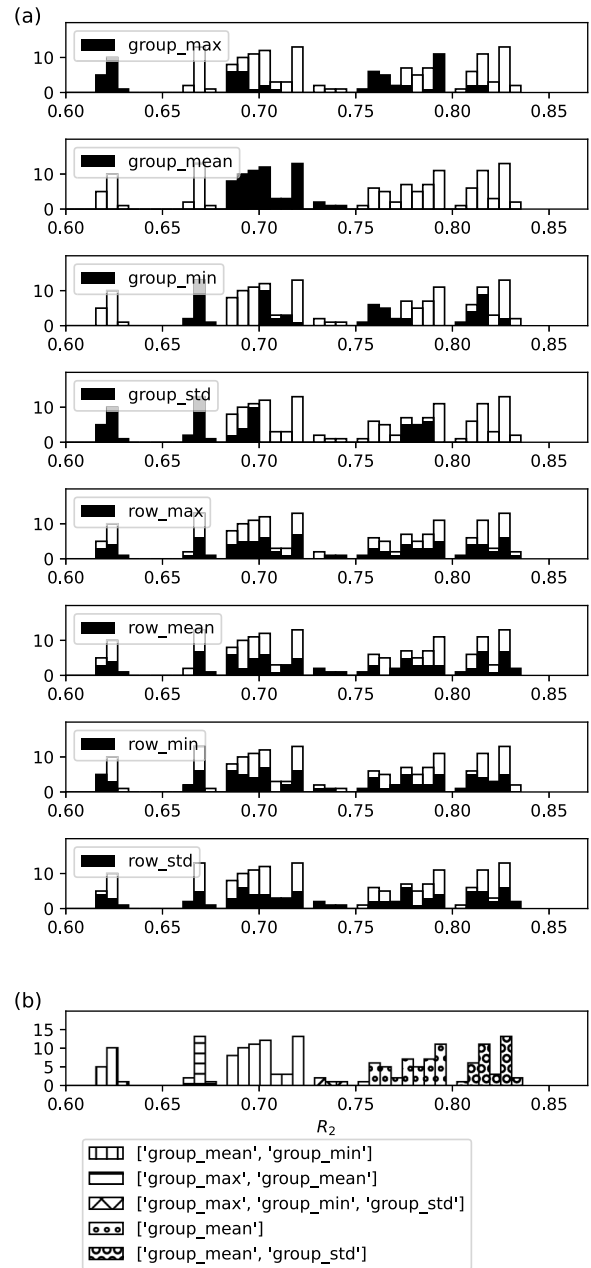


FIG. 12. Occurrence of the coefficient of determination (R^2) for the resistivity by the regression model exhaustive search is shown as histograms in (a) and (b). The abscissas in (a) and (b) are common and R^2 . (a) shows histograms of $R^2_A := \{\forall k, R^2_{\text{test}}(p_1, \dots, p_k) \mid p_1, \dots, p_k \in \mathcal{P}_k(V)\}$ and $R^2_W := \{\forall k, R^2_{\text{test}}(p_1, \dots, p_k) \mid p_1, \dots, p_k \in \mathcal{P}_k(V - V_W)\}$, where $V := \{\text{group, row}\} \times \{\text{min, max, mean, std}\}$ and $V_W := \{W \mid W \in V\}$. (b) shows the common explanatory variables identified by closed frequent itemset mining. Both figures are for the BCC phase.

variables were items, so they could be easily identified by frequent itemset mining. Figure 12(b) shows the explanatory variables that are always included (the frequency is the same as the number of transactions) in those regions by hatch. For example, the region with the highest $R^2_{\text{test}} > 0.80$ always contained the explanatory variables—group_mean and group_std. As shown in SM Table V [37], $R^2_{\text{test}} = 0.77$ used group_mean and group_std. In the region where R^2_{test} was higher than

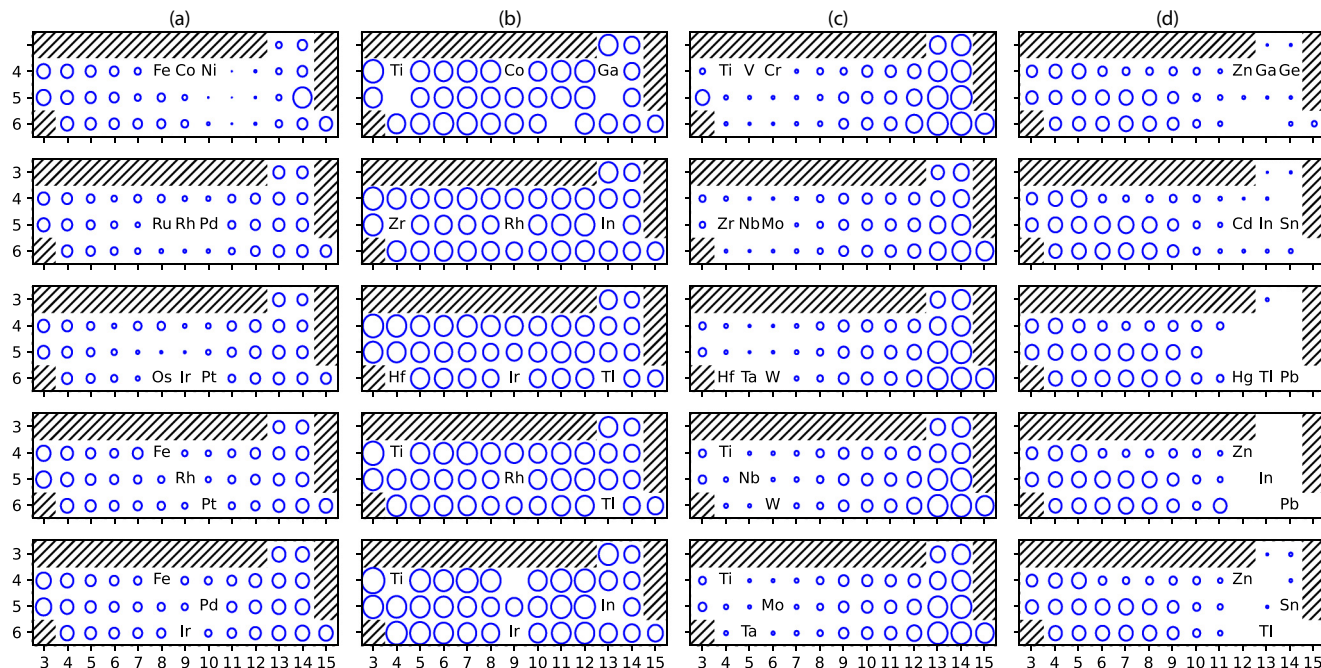


FIG. 13. Resistivities ($R(\mathcal{X}, \mathcal{Y}, \mathcal{Z}, \mathcal{W})$) are shown in the BCC phase, where $\mathcal{X}\mathcal{Y}\mathcal{Z}$ is the element indicated in the figure. The radii of the circles represent the value of R . The scale of R is the same in all figures. For example, in the upper left diagram, $\mathcal{X}\mathcal{Y}\mathcal{Z}$ are Fe, Cr, and Ni, and \mathcal{W} is the element located at the position in the diagram that mimics the periodic table. The element located at (row, column) = (4,3) in the diagram representing the periodic table is Sc, then its radius corresponds to $R(\text{Fe}, \text{Cr}, \text{Ni}, \text{Sc})$. The radius at (4,4) shows the value of $R(\text{Fe}, \text{Cr}, \text{Ni}, \text{Ti})$, (5,3) does $R(\text{Fe}, \text{Cr}, \text{Ni}, \text{V})$ and so on.

this, the model had a structure where the other explanatory variables were added to `group_mean` and `group_std`. The same results were obtained for FCC. We show them in Table VI and Fig. 14 of Ref. [37].

From the above analysis, it was expected that R varied mainly in `group_mean` and `group_std`, which were made of the columns of the periodic table of the constituent elements. To concretely confirm this hypothesis, the bubble chart (Fig. 13) shows R when elements $\mathcal{X}\mathcal{Y}\mathcal{Z}$ were fixed and \mathcal{W} was varied. The size of the ellipse was the magnitude of R , and the same minimum and maximum value of R was used for all panels. The elements $\mathcal{X}\mathcal{Y}\mathcal{Z}$ were the elements indicated in the figure, and element \mathcal{W} was the element that simulates the periodic table. In Fig. 13, the groups of $\mathcal{X}\mathcal{Y}\mathcal{Z}$ were the same for each column (a)–(d), but the combinations of rows changed for each column. Each row used row 4, row 5, row 6, and row (3,4,5), row (3,6,5) from the top as $\mathcal{X}\mathcal{Y}\mathcal{Z}$ elements. Specifically, Fig. 13(a) is FeCoNi \mathcal{W} , (b) is RuRhPd \mathcal{W} , (c) is OsIrPt \mathcal{W} , and (d) is FeIrPd \mathcal{W} , where \mathcal{W} is the element of each cell.

We consider the characteristics of each row. For example, in row (a), the R values of $\mathcal{W} = \text{Cr}, \text{Cu}$ were slightly smaller than those of other rows, but the row direction was almost unchanged, and the column direction changed. In (b), all radii appeared to be almost the same. In (c), Y had a large R ; otherwise, there was no row dependence, and R increased in the group direction away from the element used in $\mathcal{X}\mathcal{Y}\mathcal{Z}$. (d) had a small R from Cr to Co, but, except for that, there was no row dependence and R became larger in the group direction away from the elements used in $\mathcal{X}\mathcal{Y}\mathcal{Z}$. In summary, R mainly changed in response to changes in the group direction.

Next, we compare each column. In (a) and (b), the `group_mean` was almost the same, but the `group_std` differed a lot. R in (a) was smaller in the center of the group than on both sides of the group, whereas R in (b) was the same with almost no group dependence.

(a), (c), and (d) are comparisons for the cases where the `group_std` was the same but the `group_mean` was different. In this case, the value of R was smaller when the \mathcal{W} element was closer to the `group_mean`. In addition, by comparing (c) and (d), the maximum value of R in case (c) appeared to be larger than the maximum value of R in case (d). `group_max` or `group_min` needed to be added to represent this. The same analysis for the FCC phase is shown in Fig. 15 of Ref. [37].

It is also known that the addition of variable elemental additives to Cu or Al change the resistivity ratios in a similar way, mainly due to differences in the periodic table groups for the host elements in experiments [75]. The law shown in Fig. 13 is a more general one covering elements from group 3 to 14 and Bi (group 15) for HEA4, compared to the one in Ref. [75] for Cu and Al. Although much more verification is needed, this law may be generally true for the resistivity ratios of alloys, and may be useful to search for materials with the expected resistivity ratios.

G. Magnetic materials consisting of non-3d magnetic elements

There are 24 materials with relatively high T_C in the BCC phase, e.g., T_C of more than 300 K. The top 10 are listed in Table IV. Closed frequent itemset mining for $T_C \geq 300$ K showed that Bi had frequency 23, Ru had frequency 22, and the case where both Ru and Bi are included had frequency 21. The maximum T_C for the FCC phase was lower than 300 K,

TABLE IV. A list of the top 10 materials that have the BCC phase, consist of nonmagnetic elements, and have high T_C s.

Name	T_C (K)
RuTiPbBi	345
RuPdTiBi	330
RuPdHgBi	329
RuAuTiBi	326
RuPtTiBi	324
RuAgHgBi	323
RuPdPbBi	323
RuHgPbBi	322
RuAuHgBi	321
RuAgTiBi	317

and there were 20 items with $T_C \geq 200$ K. Item Ru had 19 subitems, Bi: 17, Ri_Ru: 16. The details of the FCC phase are shown in SM Table VII [37]. The magnetism of these materials could be an artifact of the large local moments of Ru and V. However, the distribution of the local moments of Ru and V showed a gentle change. It would be interesting to find materials without $3d$ transition metals and with a high T_C , although it is necessary to have a solid solution phase and uniform elemental distribution.

H. Comments related to the synthesis conditions

Finally, comments related to the synthesis conditions at high temperatures are presented. The materials in the solid solution phase are in the paramagnetic state or their magnetizations are very small. In this case, interactions between elemental atoms other than local spins and the entropy contribution at high temperatures were expected to dominate and determine the elemental distributions. Although high (middle) entropy alloys were not the most stable at room temperature, they existed as a fixed elemental configuration because the mobility of atoms was sufficiently small that atomic migration seldom occurred. Depending on the combination of atoms, magnetic solutions existed at room temperature, and magnetization was generated on each atom. If the magnetization was large, a locally stable competing solution is possible with respect to the direction of the magnetization. If the spin-spin interaction is energetically dominant at room temperature, two scenarios can be considered. One is the case where the electron kinetic energy is gained and the magnetic solution becomes ferromagnetic, which may be the case for many local magnetization directions different from those shown in Fig. 9, but further analysis involving the electronic state is needed.

The other is the case where the local spins remain antiparallel. In this study, we approximated the distribution to be uniform in real space and within a collinear spin configuration. Even in this case, the frustration of the spin configuration due to the antiferromagnetic interaction resulted in a large reduction in T_C . Similar to the case in which each element was randomly distributed in real space and the antiferro-

magnetic interaction spins maintained structural frustration, in the case of HEA, there was *interelement frustration* in the magnetization between elements. Further, if we consider that the magnetization can be in a noncollinear configuration, there is a great possibility that we can obtain nonmagnetic solutions such as spin glasses. If these phenomena exist, they can be called the cocktail effect of HEA. There may be such interesting phenomena in HEA from a fundamental physics perspective.

VII. SUMMARY

Automatic exhaustive exploration was performed for large material space consisting of 147,630 equiatomic quaternary high entropy alloys. Our high-throughput calculation tool succeeded in calculating approximately 99.1% of the material space, where the numerical parameters and convergence procedure were fully automatically controlled. Combining AkaiKKR program package and supercomputer Fugaku, we constructed the unique and useful material database, which contains the total energy, magnetization, Curie temperature, and residual resistivity, for screening new functional magnetic materials. Frequent itemset mining was used to identify the characteristics of parcels in magnetization and Curie temperature space. This method clearly specified the elemental dependence of the local spin moment and spin-parallel or antiparallel configurations. The effective additive elements to the magnetic properties were also analyzed. We showed that a valid prediction model for the residual resistivity could be realized. Furthermore, we clarified the important explanatory variables and illustrated the elementary dependence based on the periodic table. Finally, the automatic exhaustive calculations and data science approaches demonstrated in this work can be easily applied to other nonstoichiometric and substitutionally disordered systems. In particular, they will be quite useful for exploring high-performance and optimized soft magnets, permanent magnets, spintronics materials, and electric resistance materials.

ACKNOWLEDGMENTS

We thank Hiroaki Honjo for helpful discussions on wire resistance. The authors acknowledge the financial support from JST CREST (Grant Nos. JPMJCR1777 and JPMJCR1812) and MEXT KAKENHI (Grant Nos. 18K04926, 20K05068, and 21H01375). This work is also partly supported by MEXT as “Program for Promoting Researches on the Supercomputer Fugaku” (DPMSD, Project ID: JPMXP1020200307) and used computational resources of supercomputer Fugaku provided by the RIKEN Center for Computational Science through the HPCI System Research Project (Project ID: hp210179). T.F., T.C., and H.K. acknowledge the JST-Mirai Program (Grant Nos. JPMJMI18G5 and JPMJMI21G2). T.F. thanks the support from “Building of Consortia for the Development of Human Resources in Science and Technology.”

[1] G. Ceder, Y.-M. Chiang, D. R. Sadoway, M. K. Aydinol, Y.-I. Jang, and B. Huang, Identification of cathode materials for

lithium batteries guided by first-principles calculations, *Nature (London)* **392**, 694 (1998).

- [2] L. Yu and A. Zunger, Identification of Potential Photovoltaic Absorbers Based on First-Principles Spectroscopic Screening of Materials, *Phys. Rev. Lett.* **108**, 068701 (2012).
- [3] W. H. Butler, X.-G. Zhang, T. C. Schulthess, and J. M. MacLaren, Spin-dependent tunneling conductance of Fe|MgO|Fe sandwiches, *Phys. Rev. B* **63**, 054416 (2001).
- [4] K. Sato, L. Bergqvist, J. Kudrnovský, P. H. Dederichs, O. Eriksson, I. Turek, B. Sanyal, G. Bouzerar, H. Katayama-Yoshida, V. A. Dinh, T. Fukushima, H. Kizaki, and R. Zeller, First-principles theory of dilute magnetic semiconductors, *Rev. Mod. Phys.* **82**, 1633 (2010).
- [5] I. Galanakis, P. H. Dederichs, and N. Papanikolaou, Origin and properties of the gap in the half-ferromagnetic Heusler alloys, *Phys. Rev. B* **66**, 134428 (2002).
- [6] <https://nomad-coe.eu>.
- [7] J. E. Saal, S. Kirklin, M. Aykol, B. Meredig, and C. Wolverton, Materials design and discovery with high-throughput density functional theory: The Open Quantum Materials Database (OQMD), *JOM* **65**, 1501 (2013).
- [8] S. Kirklin, J. E. Saal, B. Meredig, A. Thompson, J. W. Doak, M. Aykol, S. Ruehl, and C. Wolverton, The Open Quantum Materials Database (OQMD): assessing the accuracy of DFT formation energies, *NPJ Comput. Mater* **1**, 15010 (2015).
- [9] S. Curtarolo, W. Setyawan, G. L. Hart, M. Jahnatek, R. V. Chepulskii, R. H. Taylor, S. Wang, J. Xue, K. Yang, O. Levy, M. J. Mehl, H. T. Stokes, D. O. Demchenko, and D. Morgan, AFLOW: An automatic framework for high-throughput materials discovery, *Comput. Mater. Sci.* **58**, 218 (2012).
- [10] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. A. Persson, Commentary: The Materials Project: A materials genome approach to accelerating materials innovation, *APL Mater.* **1**, 011002 (2013).
- [11] W. H. Butler, Theory of electronic transport in random alloys: Korringa-Kohn-Rostoker coherent-potential approximation, *Phys. Rev. B* **31**, 3260 (1985).
- [12] P. Hohenberg and W. Kohn, Inhomogeneous electron gas, *Phys. Rev.* **136**, B864 (1964).
- [13] W. Kohn and L. J. Sham, Self-consistent equations including exchange and correlation effects, *Phys. Rev.* **140**, A1133 (1965).
- [14] J. Korringa, On the calculation of the energy of a Bloch wave in a metal, *Physica* **13**, 392 (1947).
- [15] W. Kohn and N. Rostoker, Solution of the Schrödinger equation in periodic lattices with an application to metallic lithium, *Phys. Rev.* **94**, 1111 (1954).
- [16] H. Shiba, A reformulation of the coherent potential approximation and its applications, *Prog. Theor. Phys.* **46**, 77 (1971).
- [17] P. Soven, Application of the coherent potential approximation to a system of muffin-tin potentials, *Phys. Rev. B* **2**, 4715 (1970).
- [18] H. Akai, Fast Korringa-Kohn-Rostoker coherent potential approximation and its application to FCC Ni-Fe systems, *J. Phys.: Condens. Matter* **1**, 8045 (1989).
- [19] <http://kkri.issp.u-tokyo.ac.jp>.
- [20] A. Liechtenstein, M. Katsnelson, V. Antropov, and V. Gubanov, Local spin density functional approach to the theory of exchange interactions in ferromagnetic metals and alloys, *J. Magn. Magn. Mater.* **67**, 65 (1987).
- [21] H. Mano, A new cluster approximation to the diluted Ising and Heisenberg ferromagnets, *Prog. Theor. Phys.* **57**, 1848 (1977).
- [22] H. Mano and H. Miwa, Role of second-neighbor interaction in diluted ferromagnetic systems, *Prog. Theor. Phys.* **58**, 1 (1977).
- [23] S. Hilbert and W. Nolting, Disorder in diluted spin systems, *Phys. Rev. B* **70**, 165203 (2004).
- [24] G. Bouzerar, T. Ziman, and J. Kudrnovský, Calculating the Curie temperature reliably in diluted III-V ferromagnetic semiconductors, *Europhys. Lett.* **69**, 812 (2005).
- [25] C. Takahashi, M. Ogura, and H. Akai, First-principles calculation of the Curie temperature Slater-Pauling curve, *J. Phys.: Condens. Matter* **19**, 365233 (2007).
- [26] R. Kubo, Statistical-mechanical theory of irreversible processes. I. General theory and simple applications to magnetic and conduction problems, *J. Phys. Soc. Jpn.* **12**, 570 (1957).
- [27] D. A. Greenwood, The Boltzmann equation in the theory of electrical conduction in metals, *Proc. Phys. Soc. London* **71**, 585 (1958).
- [28] J. Banhart, Relativistic and non-relativistic electron transport in disordered alloys I. Theory, *Philos. Mag. B* **77**, 85 (1998).
- [29] T. Fukushima, H. Katayama-Yoshida, K. Sato, M. Ogura, R. Zeller, and P. H. Dederichs, Local energies and energy fluctuations – applied to the high entropy alloy CrFeCoNi, *J. Phys. Soc. Jpn.* **86**, 114704 (2017).
- [30] M. Ogura, T. Fukushima, R. Zeller, and P. H. Dederichs, Structure of the high entropy alloy AlCrFeCoNi: fcc versus bcc, *J. Alloys Compd.* **715**, 454 (2017).
- [31] AkaiKKR code contains the table of experimentally observed atomic volumes. Reader can find it in subroutine qvolum.f.
- [32] J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized Gradient Approximation Made Simple, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [33] D. D. Koelling and B. N. Harmon, A technique for relativistic spin-polarised calculations, *J. Phys. C* **10**, 3107 (1977).
- [34] T. Miyake and H. Akai, Quantum theory of rare-earth magnets, *J. Phys. Soc. Jpn.* **87**, 041009 (2018).
- [35] H. Akai and P. H. Dederichs, A simple improved iteration scheme for electronic structure calculations, *J. Phys. C* **18**, 2455 (1985).
- [36] T. Fukushima, H. Akai, T. Chikyow, and H. Kino, Automatic exhaustive calculations of large material space by Korringa-Kohn-Rostoker coherent approximation method — Applied to equiatomic quaternary high entropy alloys, <https://doi.org/10.5281/zenodo.5606502>.
- [37] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevMaterials.6.023802> for additional material data and data analysis.
- [38] D. Ma, B. Grabowski, F. Körmann, J. Neugebauer, and D. Raabe, Ab initio thermodynamics of the CoCrFeMnNi high entropy alloy: Importance of entropy contributions beyond the configurational one, *Acta Mater.* **100**, 90 (2015).
- [39] S. Huang, W. Li, X. Li, S. Schönecker, L. Bergqvist, E. Holmström, L. K. Varga, and L. Vitos, Mechanism of magnetic transition in FeCrCoNi-based high entropy alloys, *Mater. Des.* **103**, 71 (2016).
- [40] J. Šebesta, K. Carva, and D. Legut, Evolution of the Curie temperature for a substituted Cantor alloy, *Phys. Rev. B* **103**, 064407 (2021).

- [41] Y. Lederer, C. Toher, K. S. Vecchio, and S. Curtarolo, The search for high entropy alloys: A high-throughput ab-initio approach, *Acta Mater.* **159**, 364 (2018).
- [42] M. S. Lucas, D. Belyea, C. Bauer, N. Bryant, E. Michel, Z. Turgut, S. O. Leontsev, J. Horwath, S. L. Semiatin, M. E. McHenry, and C. W. Miller, Thermomagnetic analysis of fecocrxni alloys: Magnetic entropy of high entropy alloys, *J. Appl. Phys.* **113**, 17A923 (2013).
- [43] M.-H. Tsai, Three strategies for the design of advanced high-entropy alloys, *Entropy* **18**, 252 (2016).
- [44] M.-H. Tsai, R.-C. Tsai, T. Chang, and W.-F. Huang, Intermetallic phases in high-entropy alloys: Statistical analysis of their prevalence and structural inheritance, *Metals* **9**, 247 (2019).
- [45] M.-Q. Ha, N.-D. Nguyen, V.-C. Nguyen, N. Takahiro, T. Chikyow, H. Kino, T. Miyake, T. Denœux, V.-N. Huynh, and H.-C. Dam, Evidence-based recommender system and experimental validation for high entropy alloys, *Nature Computational Science* **1**, 470 (2021).
- [46] M. Ester, H. P. Kriegel, J. Sander, and X. Xiaowei, A density-based algorithm for discovering clusters in large spatial databases with noise, *KDD'96: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining* (AAAI, 1996), pp. 226–231.
- [47] H.-P. Kriegel, P. Kröger, J. Sander, and A. Zimek, Density-based clustering, *WIREs Data Mining Knowl. Discov.* **1**, 231 (2011).
- [48] T. Uno, M. Kiyomi, and H. Arimura, LCM ver. 2: Efficient mining algorithms for Frequent/Closed/Maximal itemsets, in *Conference: FIMI '04, Proceedings of the IEEE ICDM Workshop on Frequent Itemset Mining Implementations, Brighton, UK, November 1, 2004* (CEUR-WS, 2004).
- [49] T. Uno, M. Kiyomi, and H. Arimura, LCM Ver.3: Collaboration of Array, Bitmap and Prefix Tree for frequent itemset mining, in *Proceedings of the 1st International Workshop on Open Source Data Mining: Frequent Pattern Mining Implementations, OSDM '05* (Association for Computing Machinery, New York, NY, USA, 2005), p. 77–86.
- [50] T. Uno, T. Asai, H. Arimura, and Y. Uchida, LCM: An efficient algorithm for enumerating frequent closed item sets, in *Workshop on Frequent Itemset Mining Implementations (FIMI'03)* (2003).
- [51] T. Uno, T. Asai, Y. Uchida, and H. Arimura, An efficient algorithm for enumerating closed patterns in transaction databases, in *Proceedings of Discovery Science 2004, Lecture Notes in Artificial Intelligence 3245* (2004).
- [52] T. Uno, Linear time Closed itemset Miner (LCM) ver. 2, <http://research.nii.ac.jp/~uno/codes.htm>.
- [53] H. Akai, Hyperfine field calculation for various alloy systems, *Hyperfine Interact.* **68**, 3 (1991).
- [54] H. P. J. Wijn, *Magnetic Properties of Metals* (Springer-Verlag, Berlin, Heidelberg, New York, 1991).
- [55] H. Akai, M. Akai, and J. Kanamori, Electronic Structure of Impurities in Ferromagnetic Iron. II. 3d and 4d Impurities, *J. Phys. Soc. Jpn.* **54**, 4257 (1985).
- [56] R. Zeller, Local magnetic behaviour of transition-metal impurities in nickel, *J. Phys. F* **17**, 2123 (1987).
- [57] B. Drittler, N. Stefanou, S. Blügel, R. Zeller, and P. H. Dederichs, Electronic structure and magnetic properties of dilute Fe alloys with transition-metal impurities, *Phys. Rev. B* **40**, 8203 (1989).
- [58] H. Akai, Electronic structure Ni–Pd alloys calculated by the self-consistent KKR-CPA method, *J. Phys. Soc. Jpn.* **51**, 468 (1982).
- [59] M. R. Press, S. N. Mishra, S. H. Devare, and H. G. Devare, Magnetic properties of isolated Rh and Ru impurities in Pd and Pd-Fe: A theoretical study, *Phys. Rev. B* **47**, 14988 (1993).
- [60] F. Körmann, D. Ma, D. D. Belyea, M. S. Lucas, C. W. Miller, B. Grabowski, and M. H. F. Sluiter, âtreasure mapsâ for magnetic high-entropy-alloys from theory and experiment, *Appl. Phys. Lett.* **107**, 142404 (2015).
- [61] N. H. Long, M. Ogura, and H. Akai, Effects of spin-wave excitations in half-metallic materials, *Phys. Rev. B* **85**, 224437 (2012).
- [62] H. Shinya, S. Kou, T. Fukushima, A. Masago, K. Sato, H. Katayama-Yoshida, and H. Akai, First-principles calculations of finite temperature electronic structures and transport properties of Heusler alloy Co₂MnSi, *Appl. Phys. Lett.* **117**, 042402 (2020).
- [63] H. Akai, Residual resistivity of Ni-Fe, Ni-Cr and other ferromagnetic alloys, *Physica B+C* **86–88**, 539 (1977).
- [64] C. Weber, C. Yee, K. Haule, and G. Kotliar, Scaling of the transition temperature of hole-doped cuprate superconductors with the charge-transfer energy, *Europhys. Lett.* **100**, 37001 (2012).
- [65] H. Ebert, D. Ködderitzsch, and J. Minár, Calculating condensed matter properties using the KKR-Grees function method—recent developments and applications, *Rep. Prog. Phys.* **74**, 096501 (2011).
- [66] K. Nagata, J. Kitazono, S. Nakajima, S. Eifuku, R. Tamura, and M. Okada, An exhaustive search and stability of sparse estimation for feature selection problem, *IPJS Online Transactions* **8**, 25 (2015).
- [67] T. Kuwatani, K. Nagata, M. Okada, T. Watanabe, Y. Ogawa, T. Komai, and N. Tsuchiya, Machine-learning techniques for geochemical discrimination of 2011 Tohoku tsunami deposits, *Sci. Rep.* **4**, 7077 (2014).
- [68] H. Ichikawa, J. Kitazono, K. Nagata, A. Manda, K. Shimamura, R. Sakuta, M. Okada, M. K. Yamaguchi, S. Kanazawa, and R. Kakigi, Novel method to classify hemodynamic response obtained using multi-channel fNIRS measurements into two groups: exploring the combinations of channels, *Front. Hum. Neurosci.* **8**, 480 (2014).
- [69] Y. Igarashi, K. Nagata, T. Kuwatani, T. Omori, Y. Nakanishi-Ohno, and M. Okada, Three levels of data-driven science, *J. Phys.: Conf. Ser.* **699**, 012001 (2016).
- [70] L. Yu and H. Liu, Efficient feature selection via analysis of relevance and redundancy, *J. Mach. Learn. Res.* **5**, 1205 (2004).
- [71] S. Visalakshi and V. Radha, A literature review of feature selection techniques and applications review of feature selection in data mining, in *IEEE International Conference Computational Intelligence and Computing Research* (IEEE, 2014), pp. 966–971.
- [72] H. C. Dam, V. C. Nguyen, T. L. Pham, A. T. Nguyen, K. Terakura, T. Miyake, and H. Kino, Important descriptors and descriptor groups of Curie temperatures of rare-earth transition-metal binary alloys, *J. Phys. Soc. Jpn.* **87**, 113801 (2018).
- [73] H. Kino, K. Ikuse, H.-C. Dam, and S. Hamaguchi, Characterization of descriptors in machine learning for data-

- based sputtering yield prediction, *Phys. Plasmas* **28**, 013504 (2021).
- [74] H. Kino, K. Nakamura, K. Hukushima, T. Miyake, and D. H. Chi, Maximum separated distribution with high interpretability found using an exhaustive search method – application to magnetocrystalline anisotropy, *J. Phys. Soc. Jpn.* **89**, 064802 (2020).
- [75] Y. Mae, *Understanding Materials through Comparison: A Story of Copper and Aluminum*, in Japanese (Japanese Standards Association, Tokyo, 2000).