










## Quasicrystals predicted and discovered by machine learning

Chang Liu <sup>1,\*</sup>, Koichi Kitahara <sup>2,3,\*</sup>, Asuka Ishikawa <sup>4,\*</sup>, Takanobu Hiroto <sup>5</sup>, Alok Singh <sup>5</sup>, Erina Fujita <sup>5,3</sup>, Yukari Katsura <sup>5,3</sup>, Yuki Inada <sup>3</sup>, Ryuji Tamura <sup>4,†</sup>, Kaoru Kimura <sup>5,3,‡</sup> and Ryo Yoshida <sup>1,5,6,§</sup>

<sup>1</sup>The Institute of Statistical Mathematics, Research Organization of Information and Systems, Tachikawa 190-8562, Japan

<sup>2</sup>Department of Materials Science and Engineering, National Defense Academy, Yokosuka 239-8686, Japan

<sup>3</sup>Department of Advanced Materials Science, The University of Tokyo, Kashiwa 277-8561, Japan

<sup>4</sup>Department of Materials Science and Technology, Tokyo University of Science, Tokyo 125-8585, Japan

<sup>5</sup>National Institute for Materials Science, Ibaraki 305-0047, Japan

<sup>6</sup>Department of Statistical Science, The Graduate University for Advanced Studies, Tachikawa 190-8562, Japan



(Received 27 October 2022; revised 18 June 2023; accepted 24 July 2023; published 25 September 2023)

Quasicrystals represent a class of ordered materials that have diffraction symmetry forbidden in periodic crystals. Since the first discovery of quasicrystals in 1984, approximately 100 thermodynamically stable quasicrystals have been synthesized. The discovery of new quasicrystals has led to the observation of novel physical phenomena, such as robust quantum criticality, fractal superconductivity, and peculiar long-range magnetic ordering. However, the pace of discovery of new quasicrystals has significantly slowed down, which is attributed to the lack of design principles for exploring new quasicrystals. Here, we demonstrate that machine learning can greatly accelerate the process of material discovery. Our model can predict stable quasicrystalline phases with high accuracy. With this model, we discovered three stable decagonal quasicrystals through an exhaustive screening of more than 1000 ternary aluminum alloy systems.

DOI: [10.1103/PhysRevMaterials.7.093805](https://doi.org/10.1103/PhysRevMaterials.7.093805)

### I. INTRODUCTION

A large area remains unexplored in materials space, given experimentally known materials to date. To discover innovative materials from such a vast search space, machine learning has attracted tremendous attention as a key technology for exploring new frontiers in materials science. To date, the discovery of new materials using machine learning has been reported for a variety of material systems, including soft matters [1–5] and inorganic solid-state materials [6,7]. However, in quasicrystal (QC) research, the use of machine learning has not significantly progressed. In this study, we proved that machine learning can predict the chemical compositions that form stable quasicrystalline phases. To be specific, three QCs were newly synthesized based on guidance by the machine learning algorithm.

Since the metallurgist D. Shechtman and his co-workers reported the discovery of the first QC in 1984 [8], approximately 100 stable QCs have been discovered to date [9,10]. However, a vast area remains unexplored in the materials space, which

may contain semiconducting or antiferromagnetic QCs that no one has yet discovered. Even considering only ternary aluminum–transition metal (Al–TM) systems, more than 1000 systems are yet to be explored. According to our survey, the presence or absence of quasicrystalline phases is currently known for only approximately 60 systems. Although almost 40 years have passed since the first QC was discovered, the mechanism of QC formation remains largely unknown. Finding new materials from such a vast search space is considerably difficult without solid design guidelines. Therefore, efficient alternatives to time-consuming, traditional trial-and-error approaches need to be established.

The Al–Mn icosahedral QC (IQC) was the first QC discovered by Shechtman *et al.* [8] and was synthesized as a metastable phase by a rapid melt-quenching method. A few years later, thermodynamically stable IQCs with Bergman-type clusters were found first in the Al–Li–Cu system by Sainfort and Dubost [11]. Tsai and his co-workers successively discovered stable IQCs with Mackay-type clusters in Al–Cu–(Fe, Ru, Os) [12,13] and Al–Pd–(Mn, Re) [14] systems. Tsai’s group also discovered stable decagonal QCs (DQCs) in Al–Ni–(Fe, Co, Rh) [15,16], Al–Cu–Rh [17], and Al–Pd–(Ru, Os) [18]. Luo *et al.* [19] reported the first stable IQCs with Bergman-type clusters in the Zn–Mg–rare earth metal systems. Following this, Tsai’s group discovered stable Bergman-type IQCs in the Zn–Mg–(Y, Tb, Dy, Ho, Er) systems [20] and DQCs in the Zn–Mg–(Y, Dy, Ho, Er, Tm, Lu) systems [21,22]. Furthermore, Tsai’s group found the first stable binary IQC with Tsai-type clusters in Cd–Yb [23] and its ternary variants in the Cd–Mg–(Gd, Tb, Dy, Ho, Er, Tm, Yb, Lu) systems [24,25]. Subsequently, various derivatives

\*These authors contributed equally to this work.

†tamura@rs.tus.ac.jp

‡bkimura@phys.mm.t.u-tokyo.ac.jp

§yoshidar@ism.ac.jp

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.

of these IQCs and DQCs have been synthesized by several scientists and classified into the Mackay-/Bergman-/Tsai-type clusters for IQCs and two-/four-/six-/eight-layer periodicity for DQCs, thereby establishing a foundation for current quasicrystal science [10].

As described above, most QCs studied to date were discovered by Tsai and co-workers based on serendipity. They considered QCs as electronic compounds stabilized via the Hume-Rothery mechanism [26] and searched for new compounds using the electron per atom ratio ( $e/a$ ) as a guideline. As a result,  $e/a$  has been found to assume specific values for different types of IQCs [9]. However, the applicability domain of Hume-Rothery's law remains limited to a few groups of materials, and it does not hold universally for all QCs. Furthermore, no practically established design principle exists other than the  $e/a$  rule.

Aiming to accelerate the process of discovering new QCs, in a previous work [27], we developed a machine learning algorithm to predict the chemical compositions that form thermally stable QC phases. The input variable for the model is the chemical composition, and the output variable is a class label representing QCs, approximant crystals (ACs), and others including any ordinary periodic crystal phase. The model was trained using the chemical compositions of QCs, ACs, and ordinary periodic crystals that have been discovered so far. The predicted QC phases were comprehensively compared with known experimental phase diagrams in the literature. It was reported that the prediction performance reached a significantly high level. In particular, the binary classification of the combined QC/AC class versus ordinary periodic crystals was shown to be almost perfectly predictable. Furthermore, by analyzing the behaviors of the trained model, we noticed that Hume-Rothery's electron-concentration law was autonomously learned by our algorithm. In aluminum alloys, most predicted regions of QCs and ACs overlap the region of  $e/a = 1.8$ , indicating that our model recognizes the rule of thumb that was the origin of Tsai's serendipity based only on the list of materials discovered to date. Thus, our model is called  $T_{SAI}$  (artificial intelligence as an alternative to Tsai's serendipity).

This work presents a proof-of-concept study of Liu *et al.* [27]. We demonstrate the successful discovery of QCs based on machine learning. Based on the predictions of  $T_{SAI}$ , high-throughput virtual screening was performed on the entire composition space of ternary aluminum-transition metal systems. Among the 185 systems in which the presence of QCs was hypothesized, we selected Al-Ni-Os, Al-Ir-Mn, and AlIr-Fe systems, and we succeeded in finding the three DQCs.

## II. METHODS

### A. Machine learning workflow

The problem was formulated into a supervised learning framework to classify any given chemical composition into one of three different phases: QCs, ACs, and others including any ordinary periodic crystal phase [Fig. 1(a)]. In this paper, we use QC, AC, and OTHERS to denote these three different phases, respectively. Constituent elements in a given composition were characterized by 58 element features, including

the atomic number, bond radius, van der Waals radius, electronegativity, thermal conductivity, band gap, polarizability, boiling point, and melting point (see Supplemental Material Table S2 for the full list of the element features). By taking four different summary statistics (weighted mean, weighted variance, max and minpooling) for each of the 58 element features, an input composition was translated into a 232-dimensional ( $= 58 \times 4$ ) descriptor vector that represents the physicochemical features of the constituent elements (see Supplemental Material [28] for details). As a training dataset, a list of 80 thermodynamically stable QCs (QC) and 78 ACs (AC) discovered to date was compiled from the Crystallography of Quasicrystals handbook [10], and 10 000 periodic crystal chemical compositions were randomly extracted from the Materials Project database [29] as others (OTHERS). In addition, 90 compositions from failed laboratory experiments on the synthesis of QCs were added to the dataset in the class OTHERS. For each class, approximately 80% of the overall samples were randomly selected to train the model (66, 60, and 8072 for QC, AC, and OTHERS, respectively), and the masked remaining samples were used as a test set to evaluate prediction performance (14, 18, and 2018 for QC, AC, and OTHERS, respectively).

We tested various supervised learning algorithms, such as random forests and neural networks, to build predictive models to learn the mapping from the vectorized compositions to the three different phases. Because there were no significant differences in prediction performance between the various algorithms, we decided to use the random forest classifiers to build our predictive model.

Here, note that, to handle the highly unbalanced class labels, the crystal data were undersampled from the overall instances in the Materials Project database. The performance sensitivity of the classifiers when undersampling and oversampling (ADASYN [30] and Borderline-SMOTE [31]) the crystal data to various sample sizes is reported in Supplemental Material Tables S3 and S4. In summary, there were no significant differences in performance measures with respect to sample size changes in undersampling. When oversampling was applied, the performance of the classifiers was significantly lower than that of the classifiers based on undersampling.

### B. Generalization capability

In the binary classification to predict a combined QC/AC class versus OTHERS, these two phases were almost perfectly predictable as the precision and recall rates reached 0.997 and 0.999, respectively. In addition, the learned model was demonstrated to have the ability to discriminate between QCs and ACs although the classification performance degraded slightly [Fig. 1(b)]. The precision and recall were 0.722 and 0.602 for QC and 0.731 and 0.608 for AC, respectively (see Supplemental Material [28] for the definition of precision and recall).

Here, we noticed one interesting fact. By analyzing the predicted phase diagrams of Al-TM systems, it was found that Hume-Rothery's electron-concentration law was autonomously learned by our algorithm [Fig. 1(c)]. It is known that IQCs and their ACs in Al-TM systems with Mackay-type

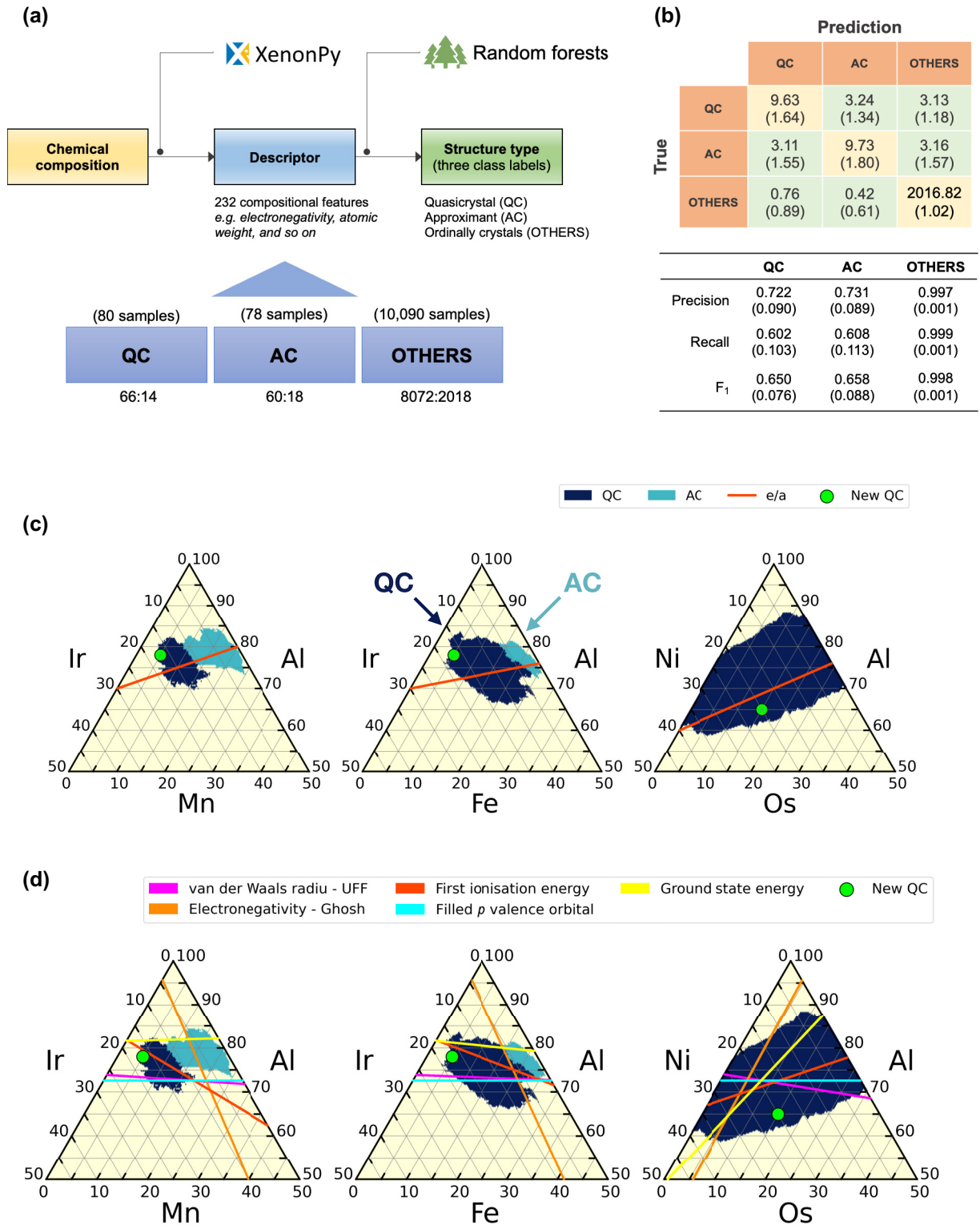


FIG. 1. Machine learning for predicting QC-forming chemical compositions. (a) Schematic of the machine learning workflow. (b) Confusion matrix (top) and three performance metrics (bottom: precision, recall, and F<sub>1</sub> scores) summarize the predictive ability of the machine learning algorithm with respect to test datasets. The mean (standard deviation) of each performance metric across 100 models trained on different datasets is shown. (c) Predicted phase diagrams of three aluminum alloys, Al-Ir-Mn, Al-Ir-Fe, and Al-NiOs. The predicted QC-/AC-forming regions are denoted by navy- and light-blue areas, respectively. Straight lines for the rule of  $e/a = 1.8$  are overlaid. Newly synthesized DQCs are colored in green. (d) Five rules for the formation of QC/AC phases, proposed by the machine learning algorithm, are represented by the straight lines on the predicted phase diagrams.

clusters tend to appear around the region of  $e/a = 1.8$  [9]. It was noticed in a *post hoc* analysis that most of the predicted regions of QCs and ACs overlap the region of  $e/a = 1.8$ , indicating that our model recognizes the rule of thumb that was the origin of Tsai's serendipity based only on the list of materials discovered to date.

In addition, we analyzed how  $T_{\text{SAI}}$  recognizes the formation and stabilization of QCs beyond the  $e/a$  rule and extracted five rules from the black-box model of  $T_{\text{SAI}}$ . The five rules are related to the weighted mean of the van der Waals radius, electronegativity, first ionization energy, number of valence electrons in the  $p$  orbitals, and energy per atom in the ground state of constituent elements. To be specific, Liu *et al.* [27] showed that in the compositional space of ternary aluminum alloys, the mapping from the compositional descriptors to the structure classes (QC, AC, and OTHERS), as determined by the black-box machine learning model, can be locally approximated by the solution of the following simultaneous equations (see Supplemental Material [28] for an explanation of the derivation process):

$$\begin{aligned}
 \sum_{i=1}^3 c^i \eta_{\text{en}}(S^i) &\approx 0.15 \pm 0.00, \\
 \sum_{i=1}^3 c^i \eta_p(S^i) &\approx 0.71 \pm 0.06, \\
 \sum_{i=1}^3 c^i \eta_{\text{vdW}}(S^i) &\approx 409.05 \pm 3.37 \text{ [pm]}, \\
 \sum_{i=1}^3 c^i \eta_{\text{ie}}(S^i) &\approx 6.49 \pm 0.09 \text{ [eV]}, \\
 \sum_{i=1}^3 c^i \eta_{\text{gs}}(S^i) &\approx -4.57 \pm 0.18 \text{ [eV]}. \quad (1)
 \end{aligned}$$

Here,  $(c^1, c^2, c^3)$  denotes the compositional ratio of the three elements ( $S^1, S^2, S^3$ ). The five element features,  $\eta_{\text{en}}, \eta_p, \eta_{\text{vdW}}, \eta_{\text{ie}},$  and  $\eta_{\text{gs}}$  represent Ghosh's electronegativity scale, the number of valence electrons in  $p$  orbitals, van der Waals radius, first ionization energy, and energy per atom in the  $T = 0$  K ground state calculated by density functional theory, respectively. The QC phase is formed in the compositional space where the weighted mean of each of the five element features is approximately equal to the value on the right-hand side of Eq. (1). These rules are represented by five straight lines on a phase diagram, as shown in Fig. 1(d). We confirmed that in aluminum alloys, most of the predicted quasicrystalline phases appear near the junction of the five straight lines.

### C. Virtual screening

With  $T_{\text{SAI}}$ , we performed a high-throughput virtual screening of 1080 systems including all of the Al-TM<sup>[4,5]</sup>-TM<sup>[4,5]</sup> and Al-TM<sup>[4,5]</sup>-TM<sup>[6]</sup> systems, where the numbers in square brackets denote the periods of TM elements. A ternary phase diagram was gridded with 20 301 points by dividing the interval of the composition ratio from 0 to 1 by 200 equally spaced

grid points. A label exhibiting the maximum probability was assigned to each grid point in the diagram. Consequently, one or more QC phases were predicted to be present in 185 systems. To further narrow the candidates, we excluded the 46 systems for which the existence of QCs or phase diagrams has been reported in the literature. Then, among the remaining systems, we selected 30 systems in which the predicted QC-phase regions are relatively large compared to other systems. As a first trial, we selected Al-Ni-Os, Al-Ir-Mn, and Al-Ir-Fe systems from the 30 systems and succeeded in finding the three DQCs.

## III. RESULTS

Figures 1(c) and 1(d) show the phase diagrams of Al-Ni-Os, Al-Ir-Mn, and Al-Ir-Fe systems predicted by  $T_{\text{SAI}}$  and the three synthesized DQCs, Al<sub>~65</sub>Ni<sub>~20</sub>Os<sub>~15</sub>, Al<sub>78</sub>Ir<sub>17</sub>Mn<sub>5</sub>, and Al<sub>78</sub>Ir<sub>17</sub>Fe<sub>5</sub>, which were successfully included in their predicted phase regions. The straight lines overlaid on the phase diagrams represent the compositional regions satisfying  $e/a = 1.8$  and the five formation rules derived from  $T_{\text{SAI}}$ . In these systems,  $T_{\text{SAI}}$  predicted the existence of QC phases near regions approximately satisfying all six rules. In terms of  $e/a$ , the predicted QC-phase regions in Al-Ni-Os, Al-Ir-Mn, and Al-Ir-Fe were distributed in the range of approximately  $e/a \in [1.5, 2.5]$ ,  $e/a \in [1.5, 2.1]$ , and  $e/a \in [1.5, 2.3]$ , respectively. Note that  $e/a = 1.8$  is the rule established for Mackay-type IQCs. As described below, the QCs discovered in these systems are not Mackay-type IQCs but DQCs. The range of  $e/a$  includes statistical uncertainty due to the inclusion of multiple types of QCs in the dataset used to train  $T_{\text{SAI}}$ , as well as to predict the value of  $e/a$  for DQCs.

We experimentally investigated the compositional regions in which the presence of QCs was predicted, and DQCs were found in the samples with the nominal compositions of Al<sub>70</sub>Ni<sub>20</sub>Os<sub>10</sub>, Al<sub>80</sub>Ir<sub>14</sub>Mn<sub>6</sub>, and Al<sub>79</sub>Ir<sub>16</sub>Fe<sub>5</sub> (see Supplemental Material [28] for experimental details). Figures 2(b)–2(d) show the electron diffraction patterns taken along the tenfold (10f) and two kinds of twofold axes (2f' and 2f), respectively, of the DQC found in Al<sub>70</sub>Ni<sub>20</sub>Os<sub>10</sub> via transmission electron microscopy (TEM). Almost perfect decagonal symmetry with a large number of sharp diffraction spots was observed in the diffraction patterns, and all spots were indexed with sextuplets of integers  $h_1 h_2 h_3 h_4 h_5 h_6$ , where  $h_i$  ( $i = 1, \dots, 5$ ) and  $h_6$  represent the components of a diffraction vector along a quasiperiodic direction (in the quasiperiodic plane) and the periodic direction (perpendicular to the quasiperiodic plane), respectively [Fig. 2(a), and see Supplemental Material for the details of the selection of basis vectors [28]]. Systematic absences of  $h_1 h_2 \bar{h}_2 \bar{h}_1 0 l$  spots in Fig. 2(d), where  $l$  is an odd integer, indicate the existence of  $c$ -glide planes perpendicular to 2f axes [10]. Note that the occurrence of 00000 $l$  spots in Fig. 2(c) is because of the multiple diffraction and, thus, does not contradict the systematic absences. In Figs. 2(c) and 2(d), intense spots among 00000 $h_6$  can be observed at 000002, where the corresponding lattice spacing  $d_{000002}$  is approximately 0.2 nm. This indicates that the DQC has a two-layer periodicity of approximately 0.4 nm along the periodic direction. Electron diffraction patterns taken from the DQCs found in Al<sub>80</sub>Ir<sub>14</sub>Mn<sub>6</sub> and Al<sub>79</sub>Ir<sub>16</sub>Fe<sub>5</sub> samples are shown in

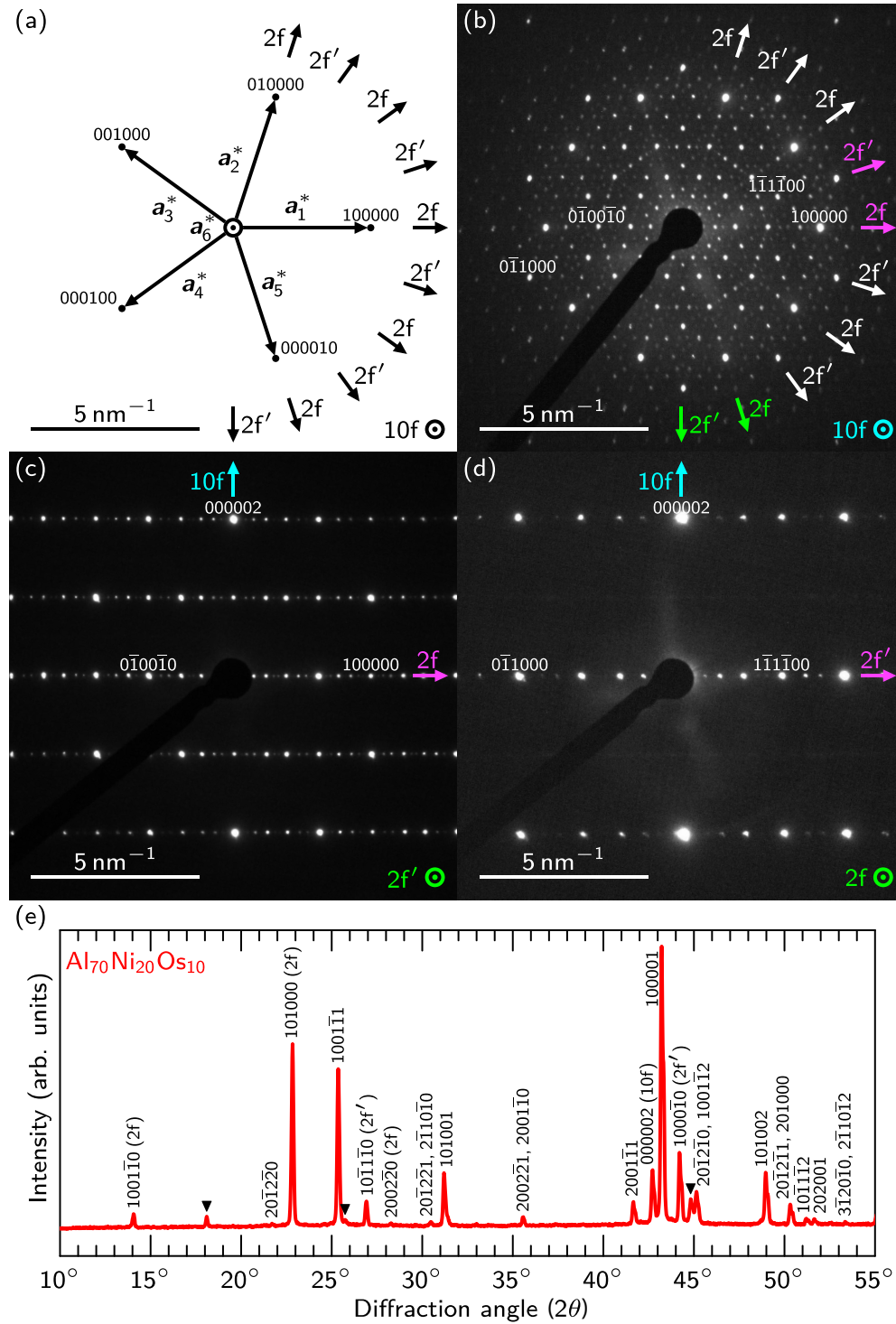


FIG. 2. (a) Basis vectors used for indexing diffraction from DQCs. Selected area electron diffraction patterns taken along (b)  $10f$ , (c)  $2f'$ , and (d)  $2f$  axes of a DQC found in an  $\text{Al}_{70}\text{Ni}_{20}\text{Os}_{10}$  sample. (e) Powder XRD pattern taken from an  $\text{Al}_{70}\text{Ni}_{20}\text{Os}_{10}$  sample with indices (and direction if the diffraction vector is along  $10f$ ,  $2f'$ , or  $2f$ ) for the DQC phase. Filled triangles indicate the diffraction from phases other than the DQC phase.

Figs. 3(a)–3(f). These patterns are interpreted in a similar way to that above, except that the DQCs found in the  $\text{Al}_{80}\text{Ir}_{14}\text{Mn}_6$  and  $\text{Al}_{79}\text{Ir}_{16}\text{Fe}_5$  samples have an eight-layer periodicity of approximately 1.6 nm along the periodic direction.

Figures 2(e) and 3(g) show powder x-ray diffraction (XRD) patterns of  $\text{Al}_{70}\text{Ni}_{20}\text{Os}_{10}$ , and  $\text{Al}_{80}\text{Ir}_{14}\text{Mn}_6$  and  $\text{Al}_{79}\text{Ir}_{16}\text{Fe}_5$

samples, respectively. For all the samples, most peaks were indexed as reflections from the DQC phases identified via TEM, and, thus, the DQC phases are the main phases in these samples. The chemical compositions of the main phases were analyzed. As shown above and in Figs. 1(c) and 1(d), these compositions are  $\text{Al}_{\sim 65}\text{Ni}_{\sim 20}\text{Os}_{\sim 15}$ ,  $\text{Al}_{78}\text{Ir}_{17}\text{Mn}_5$ , and

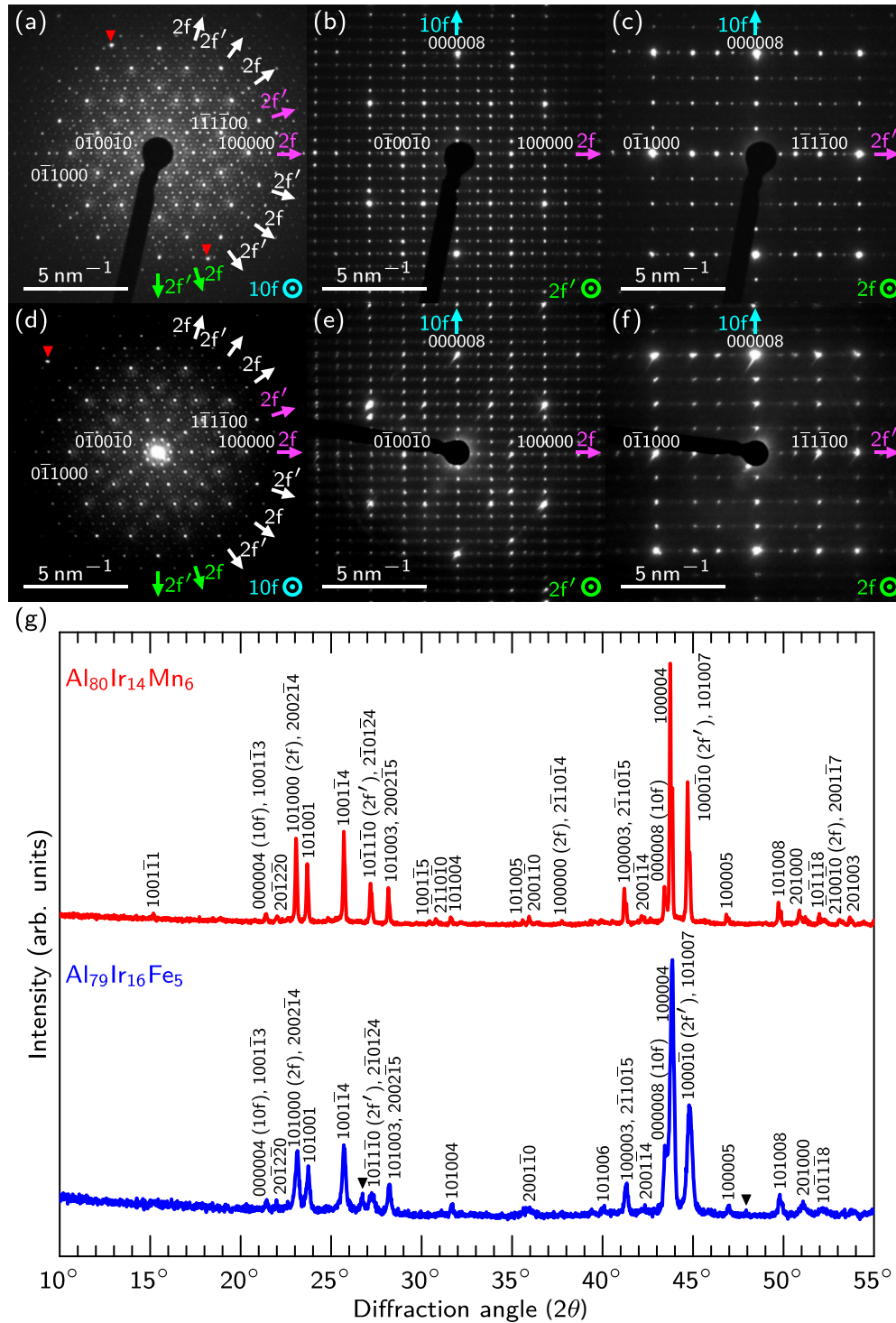


FIG. 3. Selected area electron diffraction patterns taken from (a) 10f, (b) 2f', and (c) 2f axes of a DQC found in an  $\text{Al}_{80}\text{Ir}_{14}\text{Mn}_6$  sample and those taken from (d) 10f, (e) 2f', and (f) 2f axes of a DQC found in an  $\text{Al}_{79}\text{Ir}_{16}\text{Fe}_5$  sample. (g) Powder XRD patterns taken from  $\text{Al}_{80}\text{Ir}_{14}\text{Mn}_6$  and  $\text{Al}_{79}\text{Ir}_{16}\text{Fe}_5$  samples with indices (and direction if the diffraction vector is along 10f, 2f', or 2f) for the DQC phases. Filled triangles indicate the diffraction from phases other than the DQC phases.

$\text{Al}_{78}\text{Ir}_{17}\text{Fe}_5$ . Because these DQC phases were observed after a long annealing process, they are considered to be in thermodynamically stable phases. The lattice parameters determined from the XRD peak positions and chemical compositions of the discovered DQCs are summarized in Table I, along with those of known stable DQCs with a two- or eight-layer

periodicity. Their lattice parameters are similar to those of the known DQCs. Note that the mole fraction of aluminum tends to be higher for DQCs having an eight-layer periodicity than for those having a two-layer periodicity, and this tendency also holds true for the new DQCs. Notably, the mole fraction of aluminum in  $\text{Al}_{78}\text{Ir}_{17}\text{Mn}_5$  and  $\text{Al}_{78}\text{Ir}_{17}\text{Fe}_5$  is significantly

TABLE I. Comparison of the composition (mole fractions) and (quasi)lattice parameters,  $a_R$  and  $c$ , for Al-based stable DQCs with 0.4 or 1.6 nm periodicity.

Composition	$a_R$ (nm)	$c$ (nm)	Reference
0.4 nm (two-layer) periodicity			
Al <sub>65</sub> Cu <sub>20</sub> Co <sub>15</sub>	0.2437(2)	0.41481(3)	[32,33]
Al <sub>65.0</sub> Cu <sub>14.6</sub> Co <sub>20.4</sub>	0.247(3)	0.4121(7)	[34]
Al <sub>61.9</sub> Cu <sub>18.5</sub> Rh <sub>19.6</sub>	0.2505(6)	0.4278(5)	[34]
Al <sub>57.6</sub> Cu <sub>25.9</sub> Ir <sub>16.5</sub>	0.251(5)	0.4258(5)	[34]
Al <sub>~70</sub> Ni <sub>~10</sub> Co <sub>~20</sub> <sup>a</sup>	~0.245	~0.408	[35,36]
Al <sub>70.2(3)</sub> Ni <sub>24.5(4)</sub> Fe <sub>5.3(2)</sub>	0.2450(8)	0.4105(7)	[37–39]
Al <sub>~65</sub> Ni <sub>~20</sub> Os <sub>~15</sub>	~0.252	~0.424	This work
1.6 nm (eight-layer) periodicity			
Al <sub>73</sub> Ir <sub>14.5</sub> Os <sub>12.5</sub>	0.2501(2)	1.6821(8)	[40]
Al <sub>75</sub> Pd <sub>15</sub> Os <sub>10</sub>	0.25233(2)	1.6750(3)	[40]
Al <sub>70</sub> Ni <sub>20</sub> Ru <sub>10</sub>	0.24828(4)	1.6539(3)	[40]
Al <sub>78</sub> Ir <sub>17</sub> Mn <sub>5</sub>	~0.249	~1.67	This work
Al <sub>78</sub> Ir <sub>17</sub> Fe <sub>5</sub>	~0.249	~1.66	This work

<sup>a</sup>Typical composition over a wide formation range of the DQC phase.

higher than that of the known DQCs with the same periodicity. This indicates that  $T_{SAI}$  successfully predicted new DQCs outside the compositional space in which DQCs have been discovered to date.

#### IV. CONCLUSIONS

In this study, we demonstrate that machine learning can predict new QCs. If the presence or absence of QC phase formation could be predicted based solely on the compositional patterns of previously synthesized materials, it would provide a powerful new methodology for QC research. More new materials are expected to be discovered in the future based on the machine learning predictor, including ferromagnetic [41], antiferromagnetic, and superconducting [42] QCs. To facilitate the process of validating the model and its practical use in material exploration, we have made all data and PYTHON codes publicly available.

Finally, we discuss several issues to be addressed. The use of machine learning technologies in QC research currently lags far behind in other material systems. This is mainly due to the lack of data accumulation and sharing. In particular, the lack of a comprehensive database of physical properties and structural information limits the potential applicability of machine learning in this area. In addition, most of the QCs and ACs discovered so far are binary or ternary systems. Therefore, the generalizability of the current model to multidimensional systems where less or no data are available is quite uncertain. Extending the predictability to multinary systems will be the key to discovering QCs with novel properties. Another challenge to overcome is to accelerate the process of materials synthesis. For example, for periodic systems, deep learning techniques are being developed to automate the phase identification of multiphase materials from powder x-ray diffraction patterns [43,44]. However, there has been no research on automated phase identification of quasicrystals, except for one previous study [45]. The potential of the

QC-forming composition predictor is synergistically enhanced when combined with techniques to automatically determine the presence or absence of QC phases from mass-produced multiphase samples in the laboratory. It is important to create a platform for data-driven QC research by comprehensively taking these issues into account.

Data supporting the findings of this study are available within the paper and its Supplemental Material. The digital data of chemical compositions used for machine learning are available at [46]. PYTHON codes for machine learning were implemented with the XenonPy software [47–49], which is available on GitHub [50].

#### ACKNOWLEDGMENTS

This work was supported in part by a MEXT KAKENHI Grant-in-Aid for Scientific Research on Innovative Areas (Grants No. 19H05820 and No. 19H05818) and JST CREST Grant No. JPMJCR22O3. R.Y. acknowledges financial support from Grant-in-Aid for Scientific Research (A) 19H01132 from the Japan Society for the Promotion of Science (JSPS) and JST CREST Grant No. JPMJCR19I3.

R.Y. and K. Kimura designed and conceived the project, and wrote a preliminary draft of the paper. C.L. and R.Y. designed and developed the machine learning framework and carried out the prediction of candidate compositions to be synthesized. K. Kimura and R.T. directed and supervised the experiments. E.F., Y.K., and Y.I. carried out the data extraction and curation. K. Kitahara and A.I. carried out the synthesis and characterization of the discovered QCs. T.H. and S.A. carried out the XRD and TEM measurements, respectively, and confirmed decagonal QCs. K. Kitahara, R.Y., K. Kimura, T.H., C.L., and R.T. wrote and revised the paper. All authors discussed the results and commented on the paper.

The authors declare no competing interests.

- [1] R. Gómez-Bombarelli, J. Aguilera-Iparraguirre, T. D. Hirzel, D. Duvenaud, D. Maclaurin, M. A. Blood-Forsythe, H. S. Chae, M. Einzinger, D.-G. Ha, T. Wu, G. Markopoulos, S. Jeon, H. Kang, H. Miyazaki, M. Numata, S. Kim, W. Huang, S. I. Hong, M. Baldo, R. P. Adams *et al.*, *Nat. Mater.* **15**, 1120 (2016).
- [2] S. Wu, Y. Kondo, M.-a. Kakimoto, B. Yang, H. Yamada, I. Kuwajima, G. Lambard, K. Hongo, Y. Xu, J. Shiomi, C. Schick, J. Morikawa, and R. Yoshida, *npj Comput. Mater.* **5**, 66 (2019).
- [3] J. Liang, S. Xu, L. Hu, Y. Zhao, and X. Zhu, *Mater. Chem. Front.* **5**, 3823 (2021).
- [4] K. Hatakeyama-Sato, T. Tezuka, Y. Nishikitani, H. Nishide, and K. Oyaizu, *Chem. Lett.* **48**, 130 (2019).
- [5] M. Reis, F. Gusev, N. G. Taylor, S. H. Chung, M. D. Verber, Y. Z. Lee, O. Isayev, and F. A. Leibfarth, *J. Am. Chem. Soc.* **143**, 17677 (2021).
- [6] C. Wen, Y. Zhang, C. Wang, D. Xue, Y. Bai, S. Antonov, L. Dai, T. Lookman, and Y. Su, *Acta Mater.* **170**, 109 (2019).
- [7] F. Ren, L. Ward, T. Williams, K. J. Laws, C. Wolverton, J. Hattrick-Simpers, and A. Mehta, *Sci. Adv.* **4**, eaq1566 (2018).
- [8] D. Shechtman, I. Blech, D. Gratias, and J. W. Cahn, *Phys. Rev. Lett.* **53**, 1951 (1984).
- [9] A.-P. Tsai, *Sci. Technol. Adv. Mater.* **9**, 013008 (2008).
- [10] W. Steurer and S. Deloudi, *Crystallography of Quasicrystals*, Springer Series in Materials Science Vol. 126 (Springer, Berlin, Heidelberg, 2009), p. 271.
- [11] P. Sainfort and B. Dubost, *J. Phys. Colloq.* **47**, C3-321 (1986).
- [12] A.-P. Tsai, A. Inoue, and T. Masumoto, *Jpn. J. Appl. Phys.* **26**, L1505 (1987).
- [13] A.-P. Tsai, A. Inoue, and T. Masumoto, *Jpn. J. Appl. Phys.* **27**, L1587 (1988).
- [14] A.-P. Tsai, A. Inoue, Y. Yokoyama, and T. Masumoto, *Mater. Trans., JIM* **31**, 98 (1990).
- [15] A.-P. Tsai, A. Inoue, and T. Masumoto, *Mater. Trans., JIM* **30**, 463 (1989).
- [16] A.-P. Tsai, A. Inoue, and T. Masumoto, *Philos. Mag. Lett.* **71**, 161 (1995).
- [17] A.-P. Tsai, A. Inoue, and T. Masumoto, *Mater. Trans., JIM* **30**, 666 (1989).
- [18] A.-P. Tsai, A. Inoue, and T. Masumoto, *Philos. Mag. Lett.* **64**, 163 (1991).
- [19] Z. Luo, S. Zhang, Y. Tang, and D. Zhao, *Scr. Metall. Mater.* **28**, 1513 (1993).
- [20] B. A. Niikura, A.-P. Tsai, A. Inoue, and T. Masumoto, *Philos. Mag. Lett.* **69**, 351 (1994).
- [21] T. J. Sato, E. Abe, and A.-P. Tsai, *Jpn. J. Appl. Phys.* **36**, L1038 (1997).
- [22] T. J. Sato, E. Abe, and A.-P. Tsai, *Philos. Mag. Lett.* **77**, 213 (1998).
- [23] H. Takakura, C. P. Gómez, A. Yamamoto, M. De Boissieu, and A.-P. Tsai, *Nat. Mater.* **6**, 58 (2007).
- [24] J. Q. Guo, E. Abe, and A.-P. Tsai, *MRS Proc.* **643**, 27 (2000).
- [25] J. Q. Guo, E. Abe, and A.-P. Tsai, *Phys. Rev. B* **62**, R14605 (2000).
- [26] U. Mizutani, *Hume-Rothery Rules for Structurally Complex Alloy Phases* (CRC Press, Boca Raton, 2011).
- [27] C. Liu, E. Fujita, Y. Katsura, Y. Inada, A. Ishikawa, R. Tamura, K. Kimura, and R. Yoshida, *Adv. Mater.* **33**, 2102507 (2021).
- [28] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevMaterials.7.093805> for supplemental notes on the details of machine learning workflow and experiments; supplemental tables on the quasicrystals and approximants used in the supervised learning; on element features used to calculate the compositional descriptors; and on the sensitivity of prediction accuracy when undersampling crystal data or oversampling quasicrystal and approximant instances, which includes Refs. [27,30,31,40,47–49,51–89].
- [29] 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, *APL Mater.* **1**, 011002 (2013).
- [30] H. He, Y. Bai, E. A. Garcia, and S. Li, *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)* (IEEE, Hong Kong, China, 2008), pp. 1322–1328.
- [31] H. Han, W.-Y. Wang, and B.-H. Mao, in *Advances in Intelligent Computing*, edited by D. Hutchison, T. Kanade, J. Kittler, J. M. Kleinberg, F. Mattern, J. C. Mitchell, M. Naor, O. Nierstrasz, C. Pandu Rangan, B. Steffen, M. Sudan, D. Terzopoulos, D. Tygar, M. Y. Vardi, G. Weikum, D.-S. Huang, X.-P. Zhang, and G.-B. Huang (Springer, Berlin, Heidelberg, 2005), Vol. 3644, pp. 878–887.
- [32] W. Steurer and K. H. Kuo, *Acta Crystallogr., Sect. B: Struct. Sci. Cryst. Eng. Mater.* **46**, 703 (1990).
- [33] W. Steurer and K. H. Kuo, *Philos. Mag. Lett.* **62**, 175 (1990).
- [34] P. Kuczera, J. Wolny, and W. Steurer, *Acta Crystallogr., Sect. B: Struct. Sci. Cryst. Eng. Mater.* **68**, 578 (2012).
- [35] W. Steurer, *Z. Kristallogr. - Cryst. Mater.* **219**, 391 (2004).
- [36] W. Steurer and S. Deloudi, *C. R. Phys.* **15**, 40 (2014).
- [37] I. Buganski and L. Bindi, *IUCrJ* **8**, 87 (2021).
- [38] L. Bindi, N. Yao, C. Lin, L. S. Hollister, C. L. Andronicos, V. V. Distler, M. P. Eddy, A. Kostin, V. Kryachko, G. J. MacPherson, W. M. Steinhardt, M. Yudovskaya, and P. J. Steinhardt, *Sci. Rep.* **5**, 9111 (2015).
- [39] L. Bindi, N. Yao, C. Lin, L. S. Hollister, C. L. Andronicos, V. V. Distler, M. P. Eddy, A. Kostin, V. Kryachko, G. J. MacPherson, W. M. Steinhardt, M. Yudovskaya, and P. J. Steinhardt, *Am. Mineral.* **100**, 2340 (2015).
- [40] S. Katrych, T. Weber, M. Kobas, L. Massüger, L. Palatinus, G. Chapuis, and W. Steurer, *J. Alloys Compd.* **428**, 164 (2007).
- [41] R. Tamura, A. Ishikawa, S. Suzuki, T. Kotajima, Y. Tanaka, T. Seki, N. Shibata, T. Yamada, T. Fujii, C.-W. Wang, M. Avdeev, K. Nawa, D. Okuyama, and T. J. Sato, *J. Am. Chem. Soc.* **143**, 19938 (2021).
- [42] K. Kamiya, T. Takeuchi, N. Kabeya, N. Wada, T. Ishimasa, A. Ochiai, K. Deguchi, K. Imura, and N. K. Sato, *Nat. Commun.* **9**, 154 (2018).
- [43] J.-W. Lee, W. B. Park, J. H. Lee, S. P. Singh, and K.-S. Sohn, *Nat. Commun.* **11**, 86 (2020).
- [44] J. Schuetzke, A. Benedix, R. Mikut, and M. Reischl, *IUCrJ* **8**, 408 (2021).
- [45] P. J. Lu, K. Deffeyes, P. J. Steinhardt, and N. Yao, *Phys. Rev. Lett.* **87**, 275507 (2001).
- [46] [https://raw.githubusercontent.com/yoshida-lab/XenonPy/master/samples/predict\\_hypermaterials/QC\\_AC\\_data.csv](https://raw.githubusercontent.com/yoshida-lab/XenonPy/master/samples/predict_hypermaterials/QC_AC_data.csv).
- [47] XenonPy platform, <https://github.com/yoshida-lab/XenonPy>.
- [48] H. Yamada, C. Liu, S. Wu, Y. Koyama, S. Ju, J. Shiomi, J. Morikawa, and R. Yoshida, *ACS Cent. Sci.* **5**, 1717 (2019).
- [49] S. Wu, G. Lambard, C. Liu, H. Yamada, and R. Yoshida, *Mol. Inf.* **39**, 1900107 (2020).
- [50] [https://github.com/yoshida-lab/XenonPy/tree/add\\_quasicrystal\\_prediction/samples/predict\\_hypermaterials](https://github.com/yoshida-lab/XenonPy/tree/add_quasicrystal_prediction/samples/predict_hypermaterials).



- [51] A. Seko, H. Hayashi, K. Nakayama, A. Takahashi, and I. Tanaka, *Phys. Rev. B* **95**, 144110 (2017).
- [52] L. Ward, A. Agrawal, A. Choudhary, and C. Wolverton, *npj Comput. Mater.* **2**, 16028 (2016).
- [53] K. Kitahara, Y. Takagiwa, and K. Kimura, *Mater. Trans.* **60**, 2490 (2019).
- [54] J. D. Fitz Gerald, R. L. Withers, A. M. Stewart, and A. Calka, *Philos. Mag. B* **58**, 15 (1988).
- [55] A. Singh and S. Ranganathan, *Philos. Mag. A* **74**, 821 (1996).
- [56] S. Ranganathan, E. A. Lord, N. K. Mukhopadhyay, and A. Singh, *Acta Crystallogr., Sect. A* **63**, 1 (2007).
- [57] G. Lemaître, F. Nogueira, and C. K. Aridas, *J. Mach. Learn. Res.* **18**, 1 (2017).
- [58] W. Steurer and S. Deloudi, *Crystallography of Quasicrystals: Concepts, Methods and Structures* (Springer, Berlin, Heidelberg, 2009), Vol. 126, pp. 261–271.
- [59] J. C. Slater, *J. Chem. Phys.* **41**, 3199 (1964).
- [60] mendeleev—A python resource for properties of chemical elements, ions and isotopes, ver. 0.11.0, <https://github.com/lmmmentel/mendeleev>.
- [61] M. Rahm, R. Hoffmann, and N. W. Ashcroft, *Chem. Eur. J.* **22**, 14625 (2016).
- [62] M. Rahm, R. Hoffmann, and N. W. Ashcroft, *Chem. Eur. J.* **23**, 4017 (2017).
- [63] J. Vogt and S. Alvarez, *Inorg. Chem.* **53**, 9260 (2014).
- [64] J. Meija, T. B. Coplen, M. Berglund, W. A. Brand, P. De Bièvre, M. Gröning, N. E. Holden, J. Irrgeher, R. D. Loss, T. Walczyk, and T. Prohaska, *Pure Appl. Chem.* **88**, 265 (2016).
- [65] S. P. Ong, W. D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, V. L. Chevrier, K. A. Persson, and G. Ceder, *Comput. Mater. Sci.* **68**, 314 (2013).
- [66] K. T. Tang, J. M. Norbeck, and P. R. Certain, *J. Chem. Phys.* **64**, 3063 (1976).
- [67] X. Chu and A. Dalgarno, *J. Chem. Phys.* **121**, 4083 (2004).
- [68] P. Pyykkö and M. Atsumi, *Chem. Eur. J.* **15**, 12770 (2009).
- [69] P. Pyykkö and M. Atsumi, *Chem. Eur. J.* **15**, 186 (2009).
- [70] P. Pyykkö, S. Riedel, and M. Patzschke, *Chem. Eur. J.* **11**, 3511 (2005).
- [71] P. Schwerdtfeger and J. K. Nagle, *Mol. Phys.* **117**, 1200 (2019).
- [72] W. M. Haynes, *CRC Handbook of Chemistry and Physics*, 95th ed. (CRC Press, Boca Raton, 2014).
- [73] T. Andersen, *Phys. Rep.* **394**, 157 (2004).
- [74] J. B. Mann, T. L. Meek, and L. C. Allen, *J. Am. Chem. Soc.* **122**, 2780 (2000).
- [75] J. B. Mann, T. L. Meek, E. T. Knight, J. F. Capitani, and L. C. Allen, *J. Am. Chem. Soc.* **122**, 5132 (2000).
- [76] D. C. Ghosh, *J. Theor. Comput. Chem.* **04**, 21 (2005).
- [77] J. E. Saal, S. Kirklin, M. Aykol, B. Meredig, and C. Wolverton, *JOM* **65**, 1501 (2013).
- [78] S. Kirklin, J. E. Saal, B. Meredig, A. Thompson, J. W. Doak, M. Aykol, S. Rühl, and C. Wolverton, *npj Comput. Mater.* **1**, 15010 (2015).
- [79] M. W. Gaultois, T. D. Sparks, C. K. H. Borg, R. Seshadri, W. D. Bonificio, and D. R. Clarke, *Chem. Mater.* **25**, 2911 (2013).
- [80] A. Belsky, M. Hellenbrandt, V. L. Karen, and P. Luksch, *Acta Crystallogr., Sect. B: Struct. Sci.* **58**, 364 (2002).
- [81] R. Allmann and R. Hinek, *Acta Crystallogr., Sect. A: Found. Crystallogr.* **63**, 412 (2007).
- [82] D. Zagorac, H. Muller, S. Ruehl, J. Zagorac, and S. Rehme, *J. Appl. Crystallogr.* **52**, 918 (2019).
- [83] D. G. Pettifor, *Solid State Commun.* **51**, 31 (1984).
- [84] P. Villars, K. Cenzual, J. Daams, Y. Chen, and S. Iwata, *J. Alloys Compd.* **367**, 167 (2004).
- [85] Material-agnostic platform for informatics and exploration, <https://wolverton.bitbucket.io>.
- [86] Electron configuration of the elements, <https://periodictable.com/Properties/A/ElectronConfigurationString.v.html>.
- [87] S. Alvarez, *Dalton Trans.* **42**, 8617 (2013).
- [88] N. L. Allinger, X. Zhou, and J. Bergsma, *J. Mol. Struct.: THEOCHEM* **312**, 69 (1994).
- [89] A. K. Rappe, C. J. Casewit, K. S. Colwell, W. A. Goddard, and W. M. Skiff, *J. Am. Chem. Soc.* **114**, 10024 (1992).