

# Quantum Annealing Optimization Method for the Design of Barrier Materials in Magnetic Tunnel Junctions

Kenji Nawa<sup>1,2,\*</sup>, Tsuyoshi Suzuki<sup>3</sup>, Keisuke Masuda<sup>1</sup>, Shu Tanaka<sup>4,5</sup> and Yoshio Miura<sup>1,6,†</sup>

<sup>1</sup>Research Center for Magnetic and Spintronic Materials, National Institute for Materials Science (NIMS), Sengen, Tsukuba 305-0047, Japan

<sup>2</sup>Graduate School of Engineering, Mie University, Kurima-machiya, Tsu, 514-8507, Japan

<sup>3</sup>Advanced Products Development Center, Technology and Intellectual Property HQ, TDK Corporation, Ichikawa 272-8558, Japan

<sup>4</sup>Department of Applied Physics and Physico-Informatics, Keio University, Hiyoshi, Yokohama 223-8522, Japan

<sup>5</sup>Human Biology-Microbiome-Quantum Research Center (WPI-Bio2Q), Keio University, Tokyo 108-8345, Japan

<sup>6</sup>Center for Spintronics Research Network (CSRN), Graduate School of Engineering Science, Osaka University, Machikaneyama, Toyonaka 560-8531, Japan

 (Received 2 March 2023; revised 31 May 2023; accepted 7 July 2023; published 17 August 2023)

In the field of spintronics, there is a strong demand for barrier materials in magnetic tunnel junctions (MTJs) having high tunnel magnetoresistance (TMR) and low resistance area product (RA). However, the design of such barrier materials with atomically controlled composition, vacancies, and disordering of the constituent elements is still challenging due to the combinatorial explosion of potential candidates. Very recently, a quantum annealing (QA) method has been utilized for combinatorial optimization problems in materials science. In this paper, we perform a proof-of-concept study by applying the QA approach combining with first-principles calculations and a machine-learning factorization machine (FM) to MTJs with inverse-type spinel  $\text{MgGa}_2\text{O}_4$  tunnel barrier. We treat 252 combinations of  $\text{Mg}^{2+}$  and  $\text{Ga}^{3+}$  cation disordering in the barrier layer of the MTJs and discuss the effect of the cation disordering on the structural stability, TMR, and RA. Our method is superior to simulated annealing, Bayesian optimization, and random sampling in searching for the best MTJs for low total energy and high TMR, but not so for low RA. We also revealed physical origins of high TMR and low RA behind the cation disordering. The present work highlights the applicability and advantage of materials informatics using FM + QA with first-principles calculations in designing spintronic MTJ devices.

DOI: [10.1103/PhysRevApplied.20.024044](https://doi.org/10.1103/PhysRevApplied.20.024044)

## I. INTRODUCTION

In next-generation information society, artificial intelligence (AI) will be further expanded. Accordingly, the development of AI-dedicated hardware such as logic architecture systems that can efficiently treat huge information is required [1]. Furthermore, it is important to increase processing speed by minimizing time losses due to data transport between memory and central processing unit (CPU) and reduction of the energy consumption [2,3]. To realize such technology, a key device is the nonvolatile logic system [2,3], which is the arithmetic circuit including nonvolatile storage function. As a prospective candidate of future nonvolatile memory, magnetic random-access memory (MRAM) [4] has attracted much attention because

of the nondestructive reading, long endurance, fast writing speed, scalability, and operation-voltage property [5]. Remaining problems in MRAM are higher density and lower energy consumption [6]. To this end, large output voltage and low resistance in magnetic tunnel junctions (MTJs) will be crucial for MRAM [6].

The MTJ is composed of two ferromagnetic electrodes separated by a nonmagnetic insulator, showing different resistance depending on the relative magnetization directions of the two ferromagnets. This effect is called the tunneling magnetoresistance (TMR) [7], which determines the output voltage of MRAM. To enhance the TMR, development of ferromagnetic materials such as half-metallic ferromagnets [8,9] is essential. On the other hand, the development of insulating barrier materials is also necessary to achieve low resistance around  $1 \text{ } \Omega\mu\text{m}^2$  as well as high TMR ratio of more than 300% (in optimistic definition) at room temperature. In spintronic applications, MgO has been mainly used as a barrier layer of MTJs

\*nawa@elec.mie-u.ac.jp

†Miura.Yoshio@nims.go.jp

exhibiting high TMR ratio [10–13] owing to dominant contributions from the  $\Delta_1$  coherent tunneling in MgO and the half-metallicity of the  $\Delta_1$  band of bcc-type ferromagnetic electrodes such as  $\text{Fe}_{1-x}\text{Co}_x$  [14,15]. However, in the MgO-based MTJs, it is difficult to reduce the resistance with keeping high TMR ratio, because reducing the barrier thickness and doping other elements into MgO such as Zn [16] and Ti [17] decrease the TMR as well as the resistance. Therefore, the development of alternative barrier materials is required for realizing MTJs with low resistance and high TMR ratio.

Another possible candidate of the barrier materials is spinel barrier such as  $\text{MgAl}_2\text{O}_4$  and  $\text{MgGa}_2\text{O}_4$  [18–24]. In previous works, epitaxial  $\text{Fe}/\text{MgGa}_2\text{O}_4/\text{Fe}(001)$  MTJs were fabricated experimentally, and TMR ratio up to 121% at room temperature (196% at 4K) and relatively low resistance area product (RA) were obtained [22]. Furthermore, theoretical calculations showed that an MTJ with normal-type spinel  $\text{MgGa}_2\text{O}_4$  barrier has a potential for a low resistance and high TMR ratio [24]. To obtain further high TMR ratios, the barrier height of the barrier layer in MTJs must be increased, for example, by doping Al in  $\text{MgGa}_2\text{O}_4$ . On the other hand, there is a trade-off relationship between high TMR ratio and low resistance, because MTJs with higher barrier height enhance not only TMR ratio but also resistance [25]. Thus, an atomic level control for composition of Al, Ga, Mg, and O vacancies in the barrier layer is necessary to obtain high TMR ratio and low resistance, simultaneously. However, there are too many structures to explore the appropriate barrier composition of  $\text{Mg}(\text{Ga}, \text{Al})_2\text{O}_4$  including vacancies, corresponding to  $10^{12}$  candidates even for 1–2 nm barrier thickness. It would be difficult to treat the system with such combinatorial explosions by using a machine-learning method, such as Bayesian optimization [26,27], Monte Carlo tree search [28], and genetic algorithms [29].

Recently, an exploration in the extremely wide space has been performed by quantum annealers and Ising machines [30–33], which are computers specializing in solving combinatorial optimization problems. In the quantum annealing (QA), the combinatorial optimization problem can be replaced with the Ising model. Then, the quantum annealer solves the problem by minimizing an Ising energy of the given Ising Hamiltonian, which describes the correlation between descriptors of the system expressed by the quantum bits (qubits). The application of the quantum annealer and the related systems to material informatics is also being considered. One can perform the QA by using D-Wave’s quantum annealer, where the qubit was realized by the superconducting flux ring with Josephson junctions [30]. The Ising machine can provide a good solution to solve a problem with a large number of material combinations with high speed. Kitai *et al.* proposed a method based on QA with factorization machine (FM) for design of metamaterial structures [34]. They showed that it is

possible to find a better structure in a practical time even when the system size rapidly increases. Hatakeyama *et al.* performed a comparison between the Ising machine optimization and conventional Bayesian approach for the problem of optimizing the electrolyte material, which requires a huge number of combinations ( $10^{24}$ ) to improve the conductivity of the electrolyte in a Li-ion battery [35]. They found that the Ising machine optimization is  $10^4$ – $10^8$  times faster than Bayesian optimization.

While the QA has been used in several fields of materials science, it has not yet been applied to the systems related to spintronics, such as magnetoresistive devices. Recently, Ju *et al.* explored structures of disordered spinel barrier  $\text{MgAl}_2\text{O}_4$  in MTJs based on the Bayesian optimization and first-principles calculations within a density-functional-theory (DFT) framework, and successfully clarified a physical parameter leading to the high TMR by using the linear regression method [36]. However, the effectiveness of QA for the material design of spintronic devices is still unknown. We aim to design a spinel  $\text{Mg}(\text{Ga}, \text{Al})_2\text{O}_4$  barrier that possesses high TMR and low RA, simultaneously, by controlling atomic composition, vacancy, and/or disordering. Since the number of candidate structures easily increases explosively (more than  $10^{12}$ ) as increasing the barrier thickness, materials design based on QA is strongly demanded, however, it is challenging and may not be suitable for the first stage of application. Therefore, in this study, we apply the QA for MTJs with inverse-type spinel  $\text{MgGa}_2\text{O}_4$  where cation atoms of Mg and Ga are randomly occupied at the cation site. Our proof-of-concept study covers comparison of searching performance of the QA with other methods, including simulated annealing, Bayesian optimization, and random sampling, in finding a suitable  $\text{MgGa}_2\text{O}_4$  barrier structure for low total energy, high TMR, and low RA. We also demonstrated that an analysis of the Ising model obtained during combinatorial optimization can provide a physical understanding on the properties of interest.

## II. MODEL AND METHOD

### A. MTJ structure

In spinel oxide,  $AB_2O_4$ , the anion ( $\text{O}^{2-}$ ) forms an fcc sublattice and the cations ( $A^{2+}$  and  $B^{3+}$ ) occupy tetrahedral and octahedral sites surrounded by the  $\text{O}^{2-}$  anions. Although in normal-type spinel structure the  $A^{2+}$  cation occupies the tetrahedral site and  $B^{3+}$  cation occupies the octahedral site, in inverse-type spinel structure half of  $B^{3+}$  cations occupy the tetrahedral site and the  $A^{2+}$  and the remaining  $B^{3+}$  cations occupy the octahedral site. Therefore, in this work, we focus on the disordering of  $\text{Mg}^{2+}$  and  $\text{Ga}^{3+}$  cations at the octahedral site in the inverse-type spinel  $\text{MgGa}_2\text{O}_4$  (MGO) barrier in the MTJ.

The MTJ structure was modeled by consisting of nine monolayers (MLs) of inverse-spinel MGO tunnel barrier

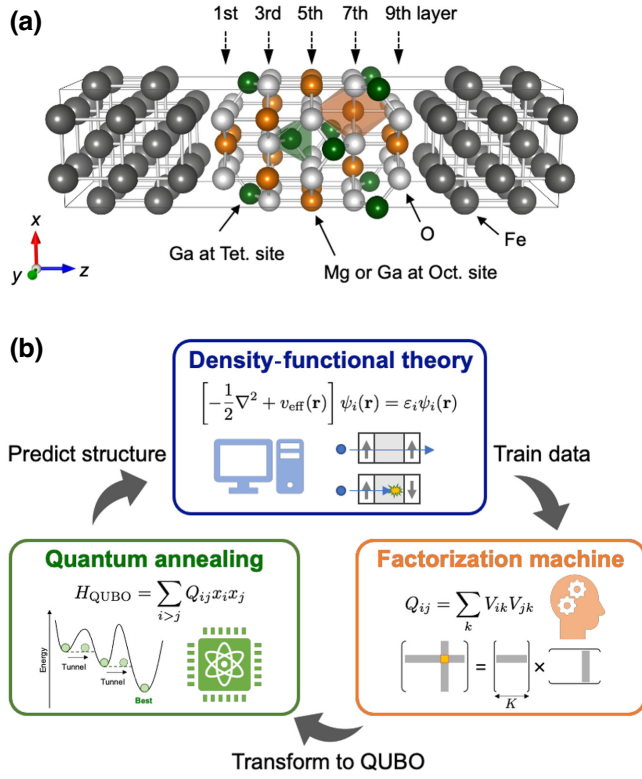


FIG. 1 (a) MTJ structure of  $\text{Fe}_{(5)}/\text{inverse-spinel MGO}_{(9)}/\text{Fe}_{(5)}$  where the number of layers is given in parenthesis in units of ML. Balls indicate Ga at tetrahedral (Tet.) site (green), disordering Mg or Ga at octahedral (Oct.) site (orange), O (white), and Fe (gray). (b) Flow of calculations combined with DFT, FM, and QA techniques.

sandwiched by 5-ML bcc Fe, as shown in Fig. 1(a). The 9-ML MGO barrier of this MTJ includes ten octahedral sites, which are randomly occupied by an equivalent number of Mg and Ga atoms; in other words, five Mg and five Ga, resulting in 252 ( $=_{10}C_5$ ) combinations of Mg-Ga occupations. The in-plane lattice constant,  $a$ , was fixed to twice that of the bulk Fe, 5.733 Å, to minimize lattice mismatch by a 45° rotation of the unit cells of Fe and MGO. This lattice constant is consistent with the experimental observations [19,20,22,37] and previous theoretical work for an MTJ with normal-type spinel MGO barrier [24]. At the Fe interface, the MGO is terminated by a layer consisting of  $\text{O}^{2-}$  anions and octahedral-site cations ( $\text{Mg}^{2+}$  or  $\text{Ga}^{3+}$ ), which is the energetically favorable geometry [24].

### B. DFT calculations

The first-principles calculations based on the DFT are performed by the linear combination of atomic orbitals (LCAO) method using the Quantum Atomistix Toolkit (Quantum-ATK) simulation [38]. The ground-state electronic structures for each MTJ are obtained within the generalized gradient approximation (GGA) [39]. The Brillouin

zone (BZ) is sampled with a Monkhorst-Pack  $k$ -point mesh of  $11 \times 11 \times 71$  along  $k_x$ ,  $k_y$ , and  $k_z$  directions, respectively. On practical calculations, an infinite open system of the Fe/MGO/Fe MTJ [referred to as the scattering region; given in Fig. 1(a)] is attached to additional 4-ML Fe electrode at right and left sides (electrode region; not shown in figure), in which a mismatch of the electronic structures between scattering and electrode regions is to be minimized by the highly dense  $k$ -point mesh along electron-transport direction,  $k_z$ . The electron wave functions are expressed by the LCAO with basis set of double- $\zeta$  polarized functions. The interfacial structure may affect the spin-dependent transport properties of the spinel-based MTJs [40,41]. Thus, atomic positions of the MGO barrier and the first interfacial Fe layer are fully relaxed by the force calculations, while all the others, i.e., the remaining 4-ML Fe in the scattering region and Fe atoms in the electrode region, are fixed to the bulk geometry. A distance of Fe/MGO interface is determined by the atomic force relaxations in each MTJ structure with the atomic force criterion of 0.05 eV/Å. Spin-transport properties are calculated using Green's function in a framework of the Landauer formula in the Quantum-ATK [38]. The number of  $\mathbf{k}_{\parallel}$  point in two-dimensional BZ is fixed to  $\mathbf{k}_{\parallel} = (k_x, k_y) = 151 \times 151$  mesh for the spin-transport calculations. This  $\mathbf{k}_{\parallel}$ -point mesh is dense enough to avoid missing a “hot-spot-like” transmittance, which may have a significant contribution to the transport property, such as interfacial resonant states [40,41].

The TMR ratio is evaluated in the pessimistic definition as  $(G_P - G_{AP})/G_P \times 100$ , where  $G_P$  and  $G_{AP}$  are conductance of parallel (P) and antiparallel (AP) magnetization states. Note that the maximum TMR in the pessimistic definition is 100%, which is different from the maximum TMR of the optimistic definition (infinity). The RA is obtained from the conductance in the P magnetization state as  $a^2/G_P$ . For a purpose of optimization with quantum annealing, we also focus on a structural stability against the disordering of Mg and Ga cation atoms. Practically, we employ a total-energy difference  $\Delta E_{\text{total}}$ , which is given as  $E_{\text{total}}^i - E_{\text{total}}^{\min}$ , where  $E_{\text{total}}^i$  is a total energy for  $i$ th MTJ structure and  $E_{\text{total}}^{\min}$  is the lowest total energy among all the MTJ structures. The total energy is taken from the P magnetization state.

### C. FM and QA

We overview the FM with QA (FM + QA) briefly, since this algorithm has been already established and described in the literature so far, for example, Refs. [32,34]. In the FM model, the target property may be given by a model equation as

$$f(\mathbf{x}) = b + \sum_{i=1}^N q_i x_i + \sum_{i=1}^{N-1} \sum_{j=i+1}^N \langle v_i, v_j \rangle x_i x_j, \quad (1)$$

where  $N$  corresponds to the number of qubits describing disordered sites ( $N = 10$  in this work) and  $\mathbf{x} = \{x_1, x_2, \dots, x_N\}$  is a descriptor vector defining cation-atom disordering of Mg and Ga. The Mg is defined as  $x_i = 0$  and the Ga is by  $x_i = 1$  for each disordered site  $i$ . Model parameters to be trained by the FM are bias term  $b$ , linear term  $q_i$ , and quadratic term  $\langle v_i, v_j \rangle$ . The  $\langle \cdot, \cdot \rangle$  is the inner product,

$$\langle v_i, v_j \rangle = \sum_{k=1}^K V_{ik} V_{jk}, \quad (2)$$

where  $K$  is a rank of FM and is treated as a hyperparameter. Use of a large  $K$  value expresses a complex FM model, while use of a small  $K$  expresses a simple FM model. In this study, the  $K$  is optimized as 7, 2, and 3 for optimizing  $\Delta E_{\text{total}}$ , TMR, and RA, respectively (see Appendix A for more details). After the FM training, the model parameters,  $q_i$  and  $\langle v_i, v_j \rangle$ , are transformed into the quadratic unconstrained binary optimization (QUBO) form to solve the Ising model via the QA. The QUBO matrix can be defined by the  $N \times N$  matrix and its matrix elements are given by

$$Q_{ij} = \begin{cases} q_i & (i = j) \\ \langle v_i, v_j \rangle & (i \neq j). \end{cases} \quad (3)$$

With this QUBO matrix, an Ising Hamiltonian we solve is written as

$$H_{\text{QUBO}} = \sum_{i < j} Q_{ij} x_i x_j + \alpha \left( \sum_{i=1}^N x_i - \frac{N}{2} \right)^2. \quad (4)$$

Here, the first term is a cost function, and the second term is a penalty term and added to fix the atomic composition of Mg and Ga, which becomes zero (nonzero) when the number of Ga atoms is (not) equivalent to that of Mg, i.e.,  $\sum_{i=1}^N x_i = 5$  ( $\sum_{i=1}^N x_i \neq 5$ ). The  $\alpha$  is a positive real-number coefficient and one of the hyperparameters in the QA to control the strength of the penalty term. The searching cycle is presented in Fig. 1(b). The next potential candidate is given by the descriptor vector  $\mathbf{x}$  that minimizes the  $H_{\text{QUBO}}$  of Eq. (4) obtained by the QA. Then, the target property is evaluated by the DFT calculations, and added as a new training data for the FM that constructs the QUBO for next cycle. This procedure is continued until the best MTJ is obtained. Thus, the combinatorial optimization problem of the inverse-spinel MGO MTJ is replaced by solving the  $H_{\text{QUBO}}$ , in which the QUBO matrix is associated with  $\Delta E_{\text{total}}$ , TMR, and RA. Since the QA algorithm seeks the minimum  $H_{\text{QUBO}}$ , in the case of TMR, maximizing the TMR is achieved by minimizing the TMR multiplied by  $-1$ , practically. Note, in our algorithm, if the next-candidate MTJ structure is already used in the

FM training, we select the next candidate randomly from the untrained candidates. We utilize the D-Wave Advantage System 4.1 as the quantum annealer through Amazon Braket Python SDK on Amazon Web Services (AWS) [42]. Our FM + QA algorithm is implemented by Python libraries of FMQA [34,43] and PyQUBO [44,45] (see Ref. [46] for more details of hyperparameters used in the FM + QA calculations).

### III. RESULTS AND DISCUSSION

#### A. DFT-calculated results

We first present results of DFT calculations for all the MTJ structures, which will be predicted by FM + QA in Sec. III B. As presented in the histogram plot of Fig. 2, the calculated values of  $\Delta E_{\text{total}}$ , TMR, and RA vary in a wide region depending on the cation distribution even at the same Mg-Ga composition (Mg : Ga = 5:5 in all the MTJs in this work). The scatter plot of Fig. 2 shows that there is no strong correlations between  $\Delta E_{\text{total}}$  and TMR [Fig. 2(a)] and between RA and  $\Delta E_{\text{total}}$  [Fig. 2(c)], whereas a correlation is visible between TMR and RA [Fig. 2(b)] (detailed discussion will be provided in Sec. III E). In the present model, the best value for each target property is referred to as the lowest  $\Delta E_{\text{total}}$  (0.0000 eV/cell), the highest TMR (60.53%), and the lowest RA (0.04826  $\Omega\mu\text{m}^2$ ), respectively, which are indicated by arrows in scatter plot of Fig. 2. From the DFT calculations, there is no inverse-spinel MGO barrier that satisfies all the preferences for low  $\Delta E_{\text{total}}$ , high TMR, and low RA in the present system.

Table I summarizes the top 6 of the optimal values for each physical property. In the  $\Delta E_{\text{total}}$ , the energy difference between top-1 and top-2 MTJs is the order of 0.01 (in eV/cell), and similarly, the difference in the TMR ratio is the order of 0.1 (in %). On the other hand, in the case of the RA, the values are approximately degenerate, with a difference of the order of  $10^{-5}$  (in  $\Omega\mu\text{m}^2$ ) between the top 1 and the top 2, and at most  $10^{-4}$  (in  $\Omega\mu\text{m}^2$ ) between the top 1 and top 6, which is several orders of magnitude smaller than those of the  $\Delta E_{\text{total}}$  and TMR. The tiny difference in the RA may bring a difficulty in predicting the best RA, as will be discussed in Sec. III B.

#### B. Prediction performance

The FM + QA results of the calculated number of structures required to obtain the best MTJ are shown in Fig. 3 for the  $\Delta E_{\text{total}}$ , TMR, and RA with the comparison of FM + simulated annealing (FM + SA), FM + exact solver (FM + ES), Bayesian optimization (BO), and random search (RS) [47]. The FM + ES is carried out as a reference, in which the Ising Hamiltonian for descriptors  $\mathbf{x}$ 's of all the training data is evaluated, from which the  $\mathbf{x}$  vector giving the lowest Ising energy of Eq. (4) is always predicted as the next candidate. The

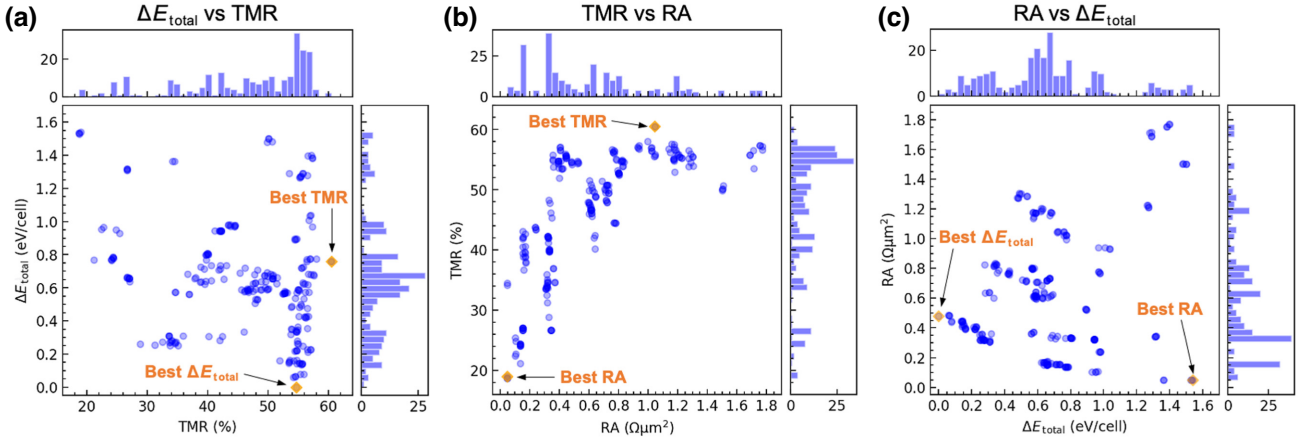


FIG. 2. DFT-calculated correlation plots for (a)  $\Delta E_{\text{total}}$  and TMR, (b) TMR and RA, and (c) RA and  $\Delta E_{\text{total}}$  with the corresponding histograms on the top and right panels. Optimal values for each property are specified by arrows.

FM + QA may require the same or a small number of calculated structures to get the best solution compared to the FM + ES, with lower computational costs than the FM + ES. Here, we performed ten trials of the FM + QA optimization using different initial training datasets containing randomly selected 20 MTJ structures, and the same datasets were also used for FM + SA, FM + ES, BO, and RS. The values of target properties were data-preprocessed by using the min-max normalization. The hyperparameters for the FM training were optimized before the search, as tabulated in Table II. The coefficient for the penalty term  $\alpha$ , the hyperparameter for the QA, is assumed to be  $\alpha = 1$ . Details in optimizations of data preprocessing and hyperparameters for FM and QA are argued in Appendix A.

For the  $\Delta E_{\text{total}}$ , the FM + QA succeeded in obtaining the best structure within around 50 structures (including training data of 20 structures) although a variation of the number of calculated structures can be seen due to the initial training dataset [Fig. 3(a)]. This result seems to be slightly superior to the FM + SA [Fig. 3(b)] and significant to BO and RS [Figs. 3(d) and 3(e)]. The efficiency of FM + QA and FM + ES [Fig. 3(c)] is comparable, meaning that our FM + QA works well in the search. For the

TMR, the performance of the FM + QA is similar with FM + SA and FM + ES, and is much better than those of the BO and RS [Figs. 3(f)–3(j)]. From a comparison of the number of calculated structures in Fig. 4, the FM + QA can be suitable in the search for the best MTJs with the lowest  $\Delta E_{\text{total}}$  and the highest TMR. The RS requires the number of calculations more than half the total searching space of 252 MTJ structures for the  $\Delta E_{\text{total}}$  and TMR.

In contrast to the case of  $\Delta E_{\text{total}}$  and TMR, the searching efficiency of the FM + QA is poor to obtain the MTJ with the lowest RA. From Fig. 3(k), there is a large variation of the number of calculated structures required to obtain the best MTJ among the ten trials. It is found that, although the FM + QA search obtains approximately the best MTJs, namely, top-1–top-6 MTJ structures, in earlier stage (approximately 25 structures), it takes many iterations to obtain the best (top-1) MTJ because of the almost degenerate RA values of top-1–top-6 structures, as mentioned in Sec. III A. The same is true for the FM + SA and FM + ES [Figs. 3(l) and 3(m)]. On the other hand, the BO needs only a few structures to find the lowest RA [Fig. 3(n)]. More insights can be obtained from Fig. 4. When the search is continued until the lowest RA is predicted (see label of RA in Fig. 4), the FM + QA, FM + SA, and FM + ES give a larger average number of calculated structures and a wider error bar than those of the BO. However, if the search is stopped once one of the approximately optimal structures (top-1–top-6) is obtained (see label of RA\* in Fig. 4), the average value of FM + QA, FM + SA, and FM + ES becomes considerably small and are superior to the RS.

From a comparison of the FM + QA and FM + SA in the RA and RA\*, we obtain the better searching performance in the FM + SA than in the FM + QA in the average value. In an energy landscape of the search space, it is known that quantum fluctuations of the QA can easily pass

TABLE I. DFT-calculated  $\Delta E_{\text{total}}$ , TMR, and RA for top-6 MTJ structures. Corresponding MTJ structures are shown in Fig. 6.

|       | $\Delta E_{\text{total}}$ (eV/cell) | TMR (%) | RA ( $\Omega\mu\text{m}^2$ ) |
|-------|-------------------------------------|---------|------------------------------|
| Top 1 | 0.0000                              | 60.53   | 0.04826                      |
| Top 2 | 0.0604                              | 60.42   | 0.04829                      |
| Top 3 | 0.0632                              | 58.01   | 0.04834                      |
| Top 4 | 0.0646                              | 57.58   | 0.04835                      |
| Top 5 | 0.0781                              | 57.57   | 0.04959                      |
| Top 6 | 0.0783                              | 57.40   | 0.04962                      |

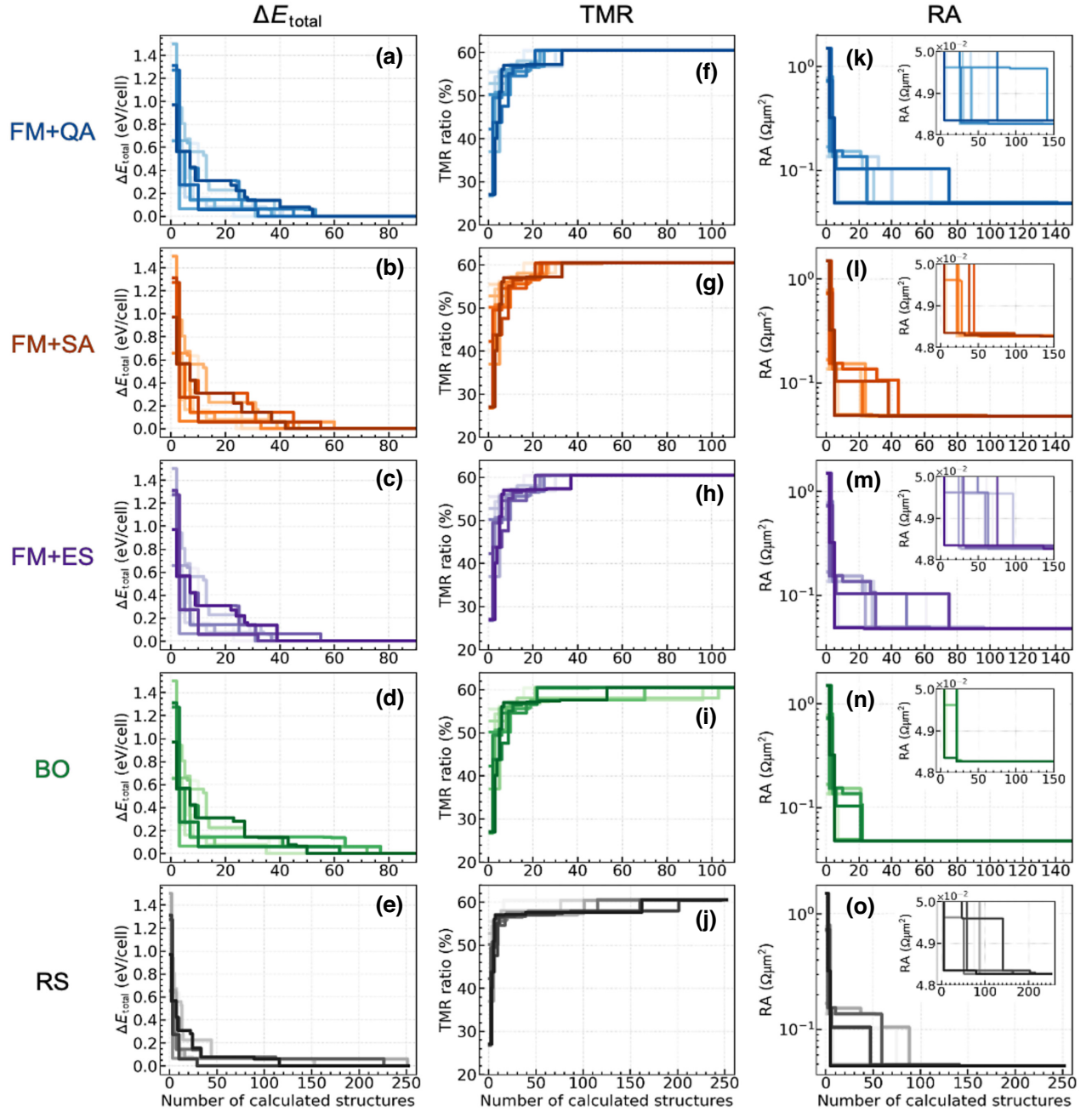


FIG. 3. History of the calculated number of structures required to obtain the best MTJ structure for (a)–(e)  $\Delta E_{\text{total}}$ , (f)–(j) TMR, and (k)–(o) RA obtained by FM + QA, FM + SA, FM + ES, BO, and RS. All the calculations are performed using a training dataset including 20 MTJ structures as the initial data, and the results of ten trials using different training datasets are plotted. The same hyperparameters, listed in Table I, are used for FM + QA, FM + SA, and FM + ES. Insets of (k)–(o) focus on the area of lower RA. The horizontal axis includes the initial training dataset, the range of which is different between the RS and others.

TABLE II. Optimized data-preprocessing and hyperparameters in FM for  $\Delta E_{\text{total}}$ , TMR, and RA.

|                           | Data preprocessing | Epoch | Rank of FM, $K$ |
|---------------------------|--------------------|-------|-----------------|
| $\Delta E_{\text{total}}$ | Min-max            | 700   | 7               |
| TMR                       | Min-max            | 100   | 2               |
| RA (in log scale)         | Min-max            | 250   | 3               |

through a narrow-energy barrier (even high barrier), but the probability of passing may be strongly suppressed for a broad-energy barrier (even low barrier), which can be easily overcome by thermal fluctuations of the SA [48]. Accordingly, we speculate that such a situation is faced in the search for the best MTJ with low RA, as illustrated in Fig. 5. Since the energy barrier separating energy

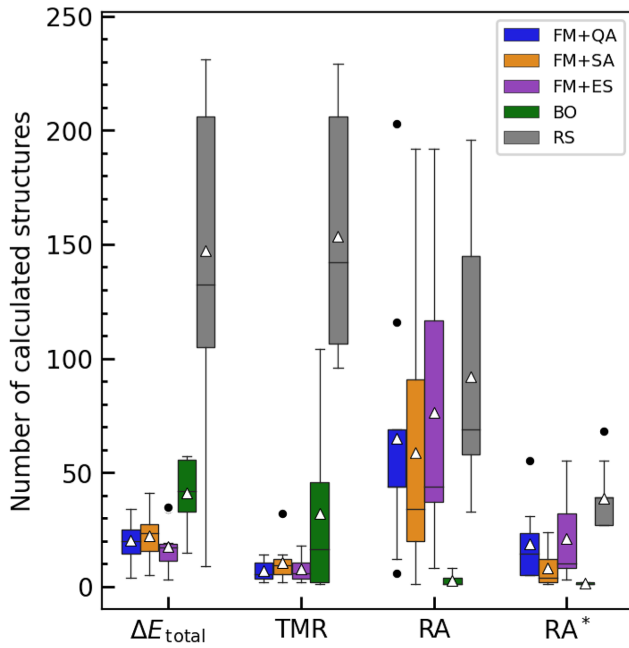


FIG. 4. Boxplot for comparison of the calculated number of structures required to obtain the best MTJ structure by FM + QA (blue), FM + SA (orange), FM + ES (purple), BO (green), and RS (gray), obtained from ten trials with different training datasets. Mean and median values are given by the triangle (white) and horizontal solid line within the box and outliers are shown in the circle (black). Initial training data (20 MTJ structures) are not included in the vertical axis. The asterisk (\*) denoted as “RA\*” indicates the results of the search until any one of top-1–top-6 structures is obtained (see the main text). All the calculations use optimized hyperparameters and  $\alpha = 1$ .

minima is the likely broad energy barrier rather than narrow barrier, the FM + QA search is less efficient than the FM + SA for the RA [see right side in Fig. 5(b)]. However, this might not be the case for the  $\Delta E_{\text{total}}$  and TMR where the FM + QA succeeds the efficient search via passing through the narrow barrier [Fig. 5(a)]. The recently proposed hybrid approach of QA and SA [49,50] may have a potential to solve such a situation, which is out of the scope of the present study.

Moreover, the poor performance of the FM + QA can be understood from the degeneracy of RA values of the DFT calculations in top-1–top-6 MTJ structures. In the energy landscape, the Ising energy minima are very close to each other [see left side in Fig. 5(b)]. This brings difficulty of constructing an Ising Hamiltonian satisfactorily to distinguish the global minimum from others in the FM, and that of finding the global energy minimum in the given energy landscape in the QA. This hypothesis might be verified from the comparison between the RA and RA\* results where the number of calculated structures becomes considerably small in the RA\*. We also mention that issues of decoherence and dissipation make

the situation more serious, which may be inevitable in the QA optimization process on a practical quantum hardware, leading to the optimized solution trapped in a local minimum [51]. One of the possible ways to overcome this difficulty is an improvement of accuracy of the FM learning process by an extremely appropriate tuning of hyperparameters, for example, by use of larger rank of FM over  $K = 10$ , although it should be aware of the overfitting in the FM training. Importantly, it notes that, the comparison of searching performance obtained in this work may not be universal since the searching efficiency of QA can be affected by various factors, including target physical properties, materials of interest, quality of constructed Ising Hamiltonian, methodology of solving the Ising Hamiltonian, and definition of descriptors.

### C. Trend of optimal MTJ structures

We discuss a trend of the optimal MTJ structures for each target property. In the present inverse-spinel MGO-based MTJ structure, there are five layers including ten octahedral cation sites randomly occupied by Mg and Ga atoms, namely, 1, 3, 5, 7, and 9-th layers, while the remaining layers (2, 4, 6, and 8-th layers) include tetrahedral cation sites occupied by Ga atom, as described in Fig. 1(a).

The top-6 optimal MTJ structures are presented in Fig. 6. For the optimal MTJs with low  $\Delta E_{\text{total}}$ , a common feature is that two octahedral cation sites at each layer contain one Mg and one Ga [see Fig. 6(a)]. This indicates that the disordering of Mg and Ga atoms favors uniform distribution in the MGO barrier, leading to the reduction of the electrostatic potential inside the barrier of MTJs. The topmost MTJ structures given in Fig. 6(a) are highly symmetric to each other; for example, the top-1 structure is identical to the top-2 by twofold rotation symmetry with respect to  $z$  axis. Other structures are also symmetric through the twofold rotation and mirror symmetry, or Mg-Ga substitution within a layer. It was confirmed that in the MTJs with higher  $\Delta E_{\text{total}}$ , the Mg and Ga cations are spatially separated, e.g., either one cation is on the left side and the other is on the right side, or either one cation is in the center of the barrier and the other is near the interface.

In the optimal MTJs with high TMR and low RA, we identified a completely opposite tendency with regard to the Mg-Ga distribution in the octahedral sites. In the MTJ with optimal TMR, Mg atoms gather around the center of MGO barrier, while Ga atoms gather near the Fe/MGO interface [see Fig. 6(b)]. We refer such a cation distribution to a Mg-center-rich distribution. In contrast, the MTJ with optimal RA is classified as a Ga-center-rich distribution, meaning that Mg atoms gather near the Fe/MGO interface and Ga atoms around the center of the barrier, respectively [see Fig. 6(c)].

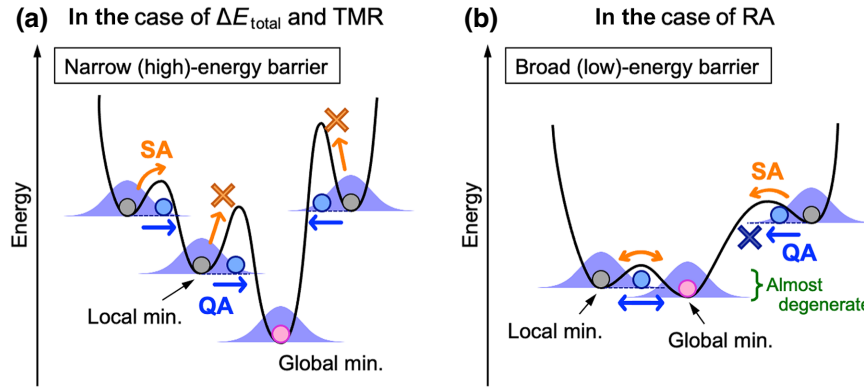


FIG. 5. Illustrations of expected energy landscape of search space and search processes of QA (blue arrow) and SA (orange arrow) in the case of (a)  $\Delta E_{\text{total}}$  and TMR and (b) RA. Global and local minima are separated by narrow (high)-energy barrier in (a) and by broad (low)-energy barrier in (b). In (b), global and local minima are almost degenerate, details of which are given in the main text.

#### D. Analysis of QUBO matrix

Interestingly, these specific cation distributions preferred for better  $\Delta E_{\text{total}}$ , TMR and RA can be addressed from the QUBO. The QUBO matrix is reconstructed by the FM model that trains all the MTJ structures, i.e., 252 MTJ structures. According to the definition of descriptor vector  $\mathbf{x}$ , the first term of Ising Hamiltonian in Eq. (4) becomes nonzero only when both the  $i$ th and  $j$ th octahedral

sites are occupied by Ga atom, i.e., the Ga-Ga case ( $\{x_i, x_j\} = \{1, 1\}$ ), otherwise zero when either site is occupied by Mg atom, i.e., the Mg-Mg case ( $\{x_i, x_j\} = \{0, 0\}$ ) or the Mg-Ga case ( $\{x_i, x_j\} = \{1, 0\}$ ). We show the QUBO matrix elements in Figs. 7(a)–7(c) and the real-space projections onto MTJ structure in Figs. 7(d)–7(f), where the red and blue colors indicate correlations between two qubits that increase or decrease the Ising energy of  $H_{\text{QUBO}}$

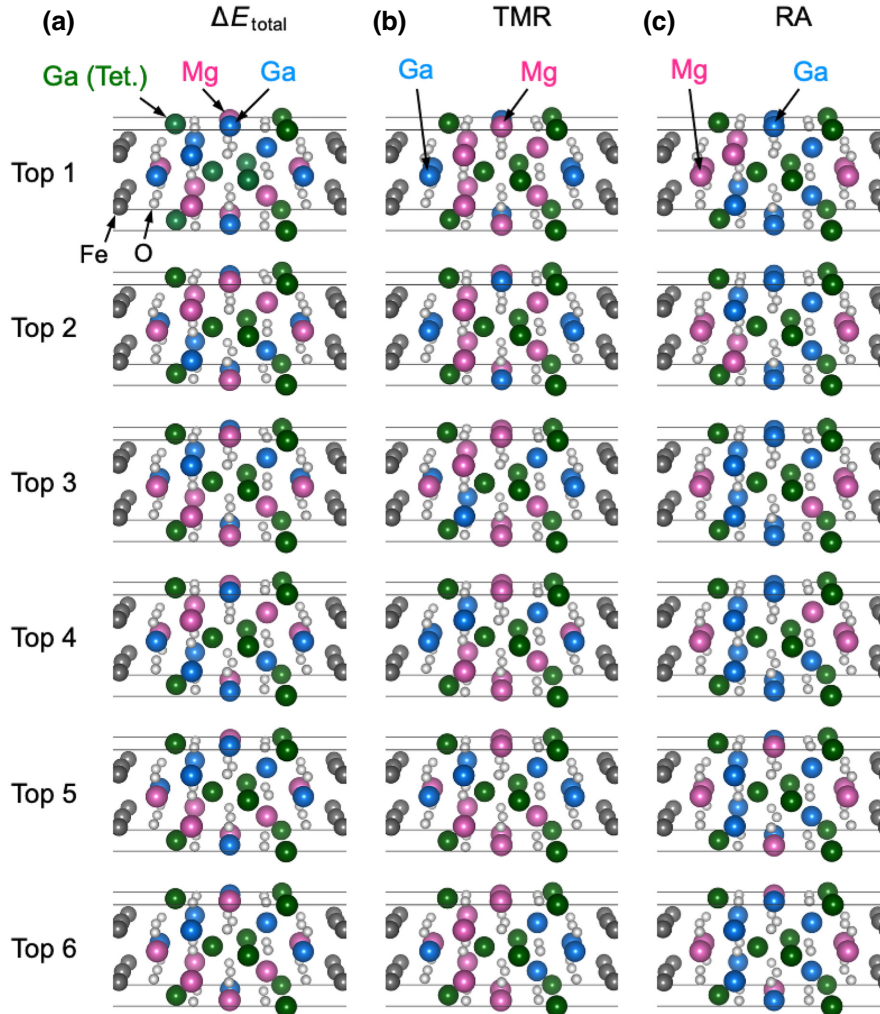


FIG. 6. The top-6 optimal MTJ structures for (a)  $\Delta E_{\text{total}}$ , (b) TMR, and (c) RA. Mg-Ga disordering cations atoms are indicated by magenta (Mg) and blue (Ga) balls, respectively.

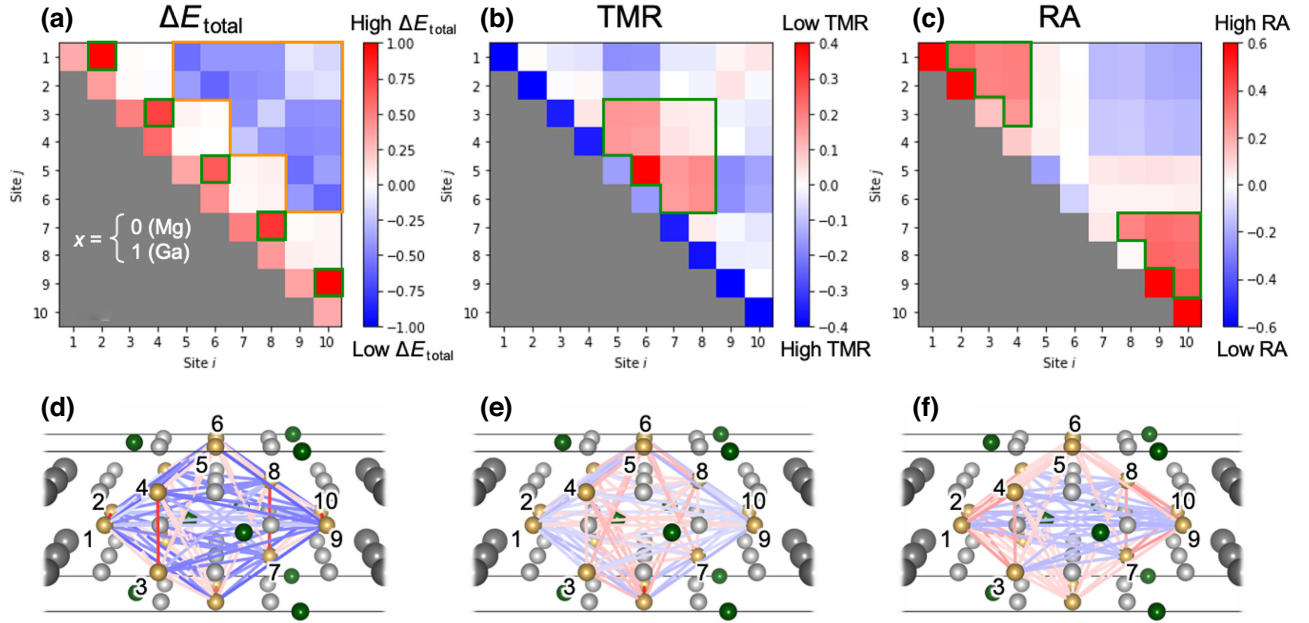


FIG. 7. (a)–(c) Matrix elements of QUBO that are reconstructed by all the 252 MTJ structures for  $\Delta E_{\text{total}}$ , TMR, and RA. Blue (red) indicates correlations of decreasing (increasing) Ising energy due to Ga-Ga intersite correlations. (d)–(f) Projections of QUBO matrix elements onto real-space MTJ structures where only the off-diagonal  $Q_{ij}$  components are projected (Ga-Ga intersite correlation). Site indexes in (a)–(c) correspond to those in (d)–(f). Note that for the TMR, the decrease (increase) of Ising energy corresponds to higher (lower) TMR.

in Eq. (4), respectively. It is noted that off-diagonal elements of QUBO matrix represent a strength of Ga-Ga correlation affecting on the target property.

Based on the above analogy, for the MTJ with low  $\Delta E_{\text{total}}$ , the off-diagonal elements next to the diagonal elements, i.e., intersite indexes of the  $i$  and  $(i+1)$ th ( $i=1, 3, 5, 7$ , and  $9$ ), show strong correlations increasing the Ising energy, as shown with red cells inside green squares in Fig. 7(a). The corresponding real-space projections are also shown by red solid lines in Fig. 7(d). This indicates that the occupation by two Ga cations at the same layer is not favorable for lower  $\Delta E_{\text{total}}$ , and two octahedral sites at the same layer should be occupied by one Ga and one Mg. On the other hand, correlations decreasing the Ising energy appear in long-distance Ga-Ga cases as observed by blue cells inside orange lines in Fig. 7(a) and blue solid lines in Fig. 7(d). This tendency of the QUBO matrix elements' distribution is consistent with the MTJ structures given in Fig. 6(a).

The QUBO matrix for the MTJ with high TMR represents that the Ga-Ga correlations in the center region of the barrier increase the Ising energy (corresponding to lowering the TMR) [see red cells inside the green line in Fig. 7(b) and red solid lines in Fig. 7(e)], resulting in the Mg-center-rich distribution. Contrarily, the MTJ with low RA shows the QUBO matrix elements that indicate the decreasing correlations near the Fe/MGO interface [red cells inside green lines in Fig. 7(c) and red solid

lines in Fig. 7(f)], resulting in the Ga-center-rich distribution. These results are also consistent with the optimal MTJ structures with high TMR and low RA shown in Figs. 6(b) and 6(c), respectively. Therefore, analysis of the QUBO matrix will be effective in understanding key correlations that determine the optimal solution of Ising Hamiltonian. In the present study, we used all the MTJ structures for the analysis of the QUBO matrix in Fig. 7, because we obtained all the data for the purpose of demonstrating the present algorithm. In general, this QUBO analysis is available for arbitrary inferred candidates during the optimization process without exploring the entire search space, once the FM + QA optimization yields the optimal solution and inputs the candidates into the FM model.

### E. Physical origins of high TMR and low RA

Finally, we discuss physical origins behind the trade-off relationship between high TMR and low RA, which are shown by the QUBO analysis. Taking the best MTJ structures, namely, the MTJs with the highest TMR and the lowest RA, we show real-space projection of local density of states (LDOS) along  $z$  axis in Figs. 8(a) and 8(b). In the Fe region, we can see metallic electronic structures characterized by the finite LDOS lying at the Fermi energy of the MTJs, while there is no LDOS in the MGO barrier region because of the energy band gap. We can estimate

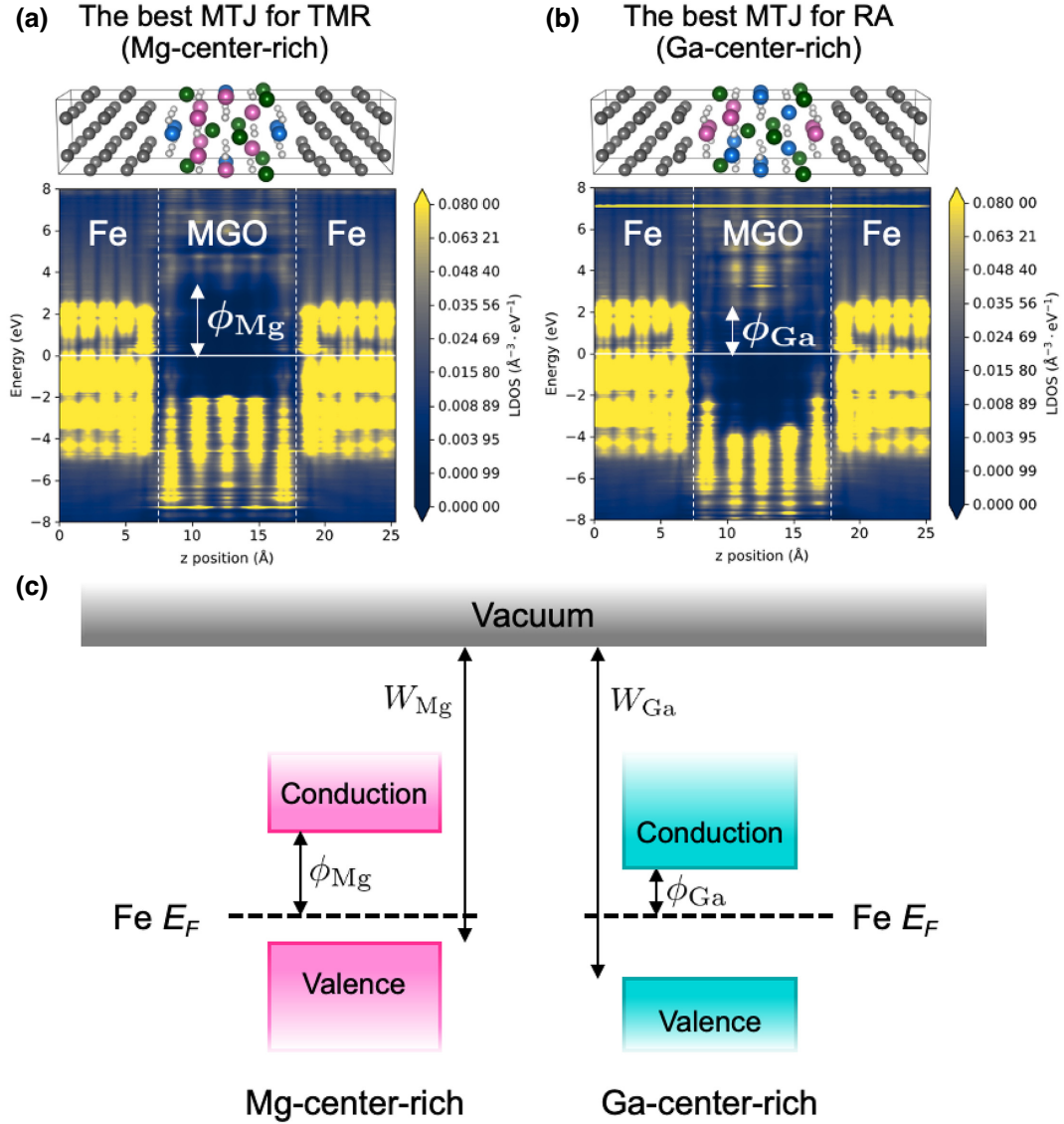


FIG. 8. Local density of states (LDOS) projected onto real-space MTJ structures for the best MTJs for (a) TMR and (b) RA. Vertical axis corresponds to energy level where Fermi energy of the MTJs is set to zero and horizontal axis corresponds to  $z$  axis in MTJ. LDOS is given as a sum of majority- and minority-spin states and is specified by the yellow colormap. (c) Illustration of energy diagram showing electronic structures of MTJs with Mg-center-rich and Ga-center-rich distributions, in which work functions and barrier heights are given by  $W_X$  and  $\phi_X$ , where  $X = \text{Mg}$  for the former and  $X = \text{Ga}$  for the latter, respectively.

the effective barrier height  $\phi$  from the energy difference between the conduction-band minimum and the Fermi energy, which are different between the Mg-center-rich distribution ( $\phi_{\text{Mg}} \sim 4$  eV) and the Ga-center-rich distribution ( $\phi_{\text{Ga}} \sim 2$  eV) [Figs. 8(a) and 8(b)]. According to the previous theoretical works [14,25], MTJs with higher barrier height provide higher TMR because the difference of the coherent tunnel conductance between P and AP magnetizations is enhanced significantly in MTJs. On the other hand, lower barrier height is responsible for low RA. Similar discussions hold for our system. The barrier height of the Mg-center-rich distribution ( $\phi_{\text{Mg}} \sim 4$  eV) is higher

than that of the Ga-center-rich distribution ( $\phi_{\text{Ga}} \sim 2$  eV), providing relatively higher TMR (60.53%) in the former and smaller RA ( $0.04826 \Omega\mu\text{m}^2$ ) in the latter. As a reference, the barrier height of the Fe/normal-spinel  $\text{MgAl}_2\text{O}_4/\text{Fe}$  MTJ showing  $\text{TMR} = 68.71\%$  is around 5 eV, which is almost comparable to that of the MGO MTJ with the Mg-center-rich distribution.

The different barrier heights between the Mg-center-rich and Ga-center-rich distributions are attributed to the difference in their work functions, not the difference in the energy band gap (see Appendix B for details in evaluating the work functions). We calculated the work functions

of the both systems and found that the work function is larger in the Ga-center-rich distribution ( $W_{\text{Ga}}=4.6$  eV) than in the Mg-center-rich distribution ( $W_{\text{Mg}}=3.8$  eV). The energy diagram of electronic structures for these MTJs is illustrated in Fig. 8(c). The larger work function indicates the deeper valence-band maximum. Therefore, the Ga-center-rich distribution has a deeper valence-band maximum than the Mg-center-rich distribution, leading to the lower barrier height for the same Fermi energy and band gap.

#### IV. SUMMARY

In summary, we conducted a proof-of-concept study to demonstrate the effectiveness of quantum annealing for the material design of spintronic devices. We considered the Fe/inverse-spinel MGO/Fe MTJs, in which the Mg and Ga atoms are disordering, with desirable physical properties such as total energy, TMR, and RA using DFT calculations combined with machine-learning factorization machine and quantum annealing, FM + QA. The results indicate the searching efficiency of FM + QA is superior to the other methods such as FM + SA, BO, and RS for the total energy and TMR, although it is insufficient compared to the BO for the RA. This insufficiency could be improved by more appropriate tuning of the hyperparameters for FM and QA routines. The QUBO matrix reconstructed by the FM during inference process provides insights into the key intersite correlation for the best physical properties. According to the QUBO-based analysis, moreover, the trade-off relationship was found between the TMR and RA in terms of the Mg-Ga cation distribution in the MGO-based MTJs. The Mg-center-rich distribution tends to prefer high TMR due to high barrier height, while the Ga-center-rich distribution tends to prefer low RA due to low barrier height. This study highlights the applicability of the quantum annealing approach in accelerating the design of spintronic tunnel barrier materials under combinatorial explosions. To demonstrate the veritable advantages of QA, further attempts are strongly needed not only for the optimization of cation-site occupation, but also for various degrees of freedom including constituent elements, atomic compositions, and atomic defects, whether ordered or disordered systems, which remain as future challenges.

#### ACKNOWLEDGMENTS

K.N. and T.S. contributed equally to this paper. Authors are grateful to S. Kawabata at National Institute of Advanced Industrial Science and Technology, H. Sukegawa and S. Mitani at NIMS, and K. Nakada at TDK Corporation for fruitful suggestions and discussions. This work is partly supported by KAKENHI (Grants No. JP20H02190, No. JP20K14782, No. JP21H01750, No. JP21K03391, No. JP22H04966, No. JP22K14290, No. JP23H05447, and No. JP23K03933);

JST (Grant No. JPMJPF2221); the TDK Corporation; and the Data-Science Research Center for Material, Quantum, and Measurement Technologies, Mie University. Human Biology-Microbiome-Quantum Research Center (Bio2Q) is supported by World Premier International Research Center Initiative (WPI), MEXT, Japan. A part of the computations was performed on the D-Wave Advantage System 4.1 through AWS and the facility of Numerical Materials Simulator at NIMS. All the data that support the findings of this study are available from the corresponding author upon reasonable request.

#### APPENDIX A: HYPERPARAMETER TUNING

Before performing the FM + QA for the prediction keeping high accuracy and efficiency, it is important to define appropriately the hyperparameters for both the FM and QA routines. For the FM, the hyperparameters include the data preprocessing for the target quantity, number of epochs for training the FM model, and the rank of FM,  $K$ , in Eq. (2).

As the data preprocessing, we consider two treatments that are mostly used in the machine learning; the standardization  $y'_i = y_i - \bar{y}/\sigma$ , where  $y_i$  is  $i$ th value of the target property  $\{y\}$  and  $\sigma$  and  $\bar{y}$  are mean and standard deviation values of  $\{y\}$ , and min-max normalization  $y'_i = (y_i - \min(y))/(\max(y) - \min(y))$ , where  $\min(y)$  and  $\max(y)$  are minimum and maximum values of  $\{y\}$ . By adapting these treatments, as well as the case without any data preprocessing, training error of the FM was evaluated for the target properties of the  $\Delta E_{\text{total}}$ , TMR, and RA. The training error is defined as a root-mean-squared error (RMSE) evaluated by fivefold cross-validation manner using the training dataset. Figures 9(a)–9(c) summarize the RMSE as a function of the number of epochs for the FM training in each property. We find that the min-max normalization obtains the smallest RMSE compared to the standardization (Std.) and no data preprocessing (None) for all the properties (see blue in figures). We emphasize that, for the results of RA, the RMSE are also compared in the log scale, in which the RA is preprocessed after logarithmic transformation, and the resultant RMSE becomes smaller than that in the linear scale [see sky blue in Fig. 9(c)]. The number of epochs for the FM training is optimized as 700, 100, and 250 from the RMSE curves of  $\Delta E_{\text{total}}$ , TMR, and RA (in log scale), respectively. With these optimized epochs, the RMSE dependence of the rank of FM ( $K$ ) is further investigated. From Figs. 9(d)–9(f), a large change of the RMSE is not found in the range of  $K = 1$ –10 that we examined in this work, except for  $\Delta E_{\text{total}}$  where use of  $K = 1$  increases the RMSE. We select the  $K$  exhibiting the smallest RMSE in this range as the optimized value, i.e.,  $K = 7, 2, 3$  for  $\Delta E_{\text{total}}$ , TMR, and RA (in log scale).

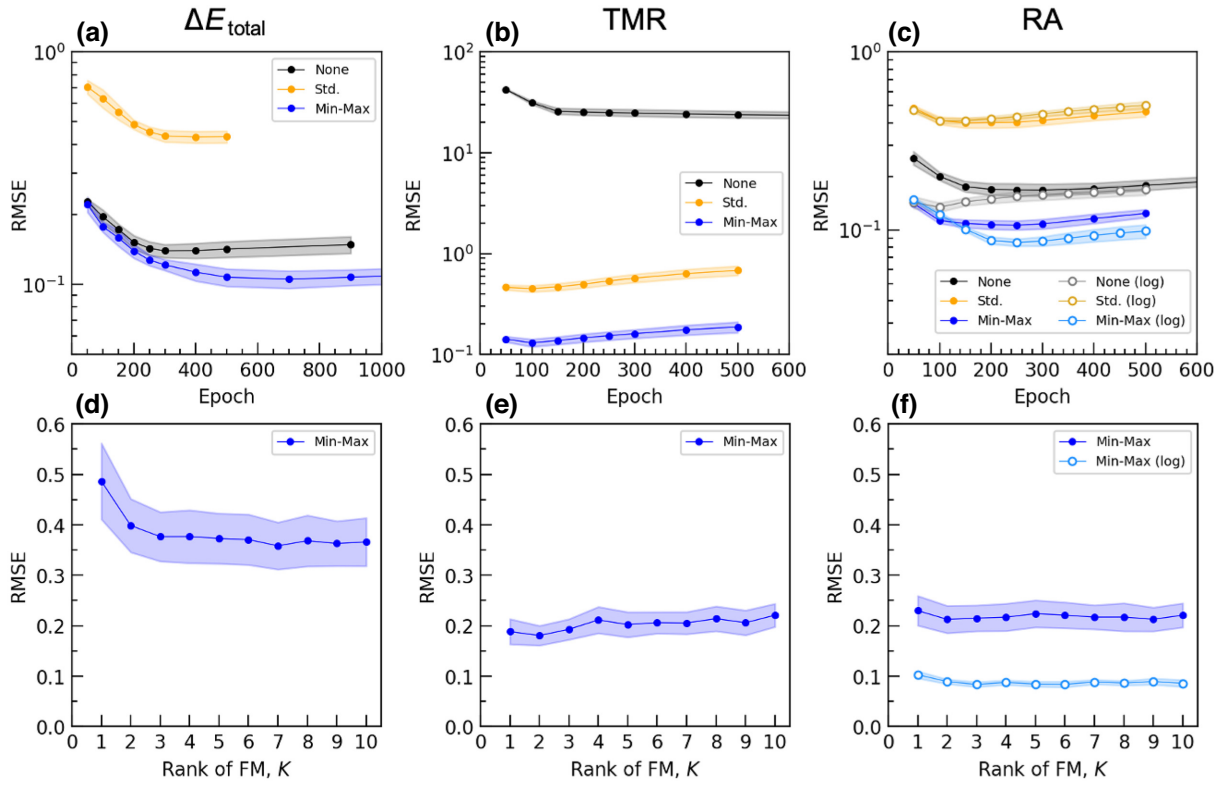


FIG. 9. (a)–(c) RMSE as a function of number of epochs in FM training, where data-preprocessing treatments of standardization (Std; orange plot) and min-max normalization (min-max; orange) are employed, for  $\Delta E_{\text{total}}$ , TMR, and RA. Results without data preprocessing is also shown (None; black) for a reference. (d)–(f) RMSE as a function of rank of FM,  $K$ , obtained by min-max normalization. In (c),(f), results of RA in log scale are also plotted by opened circles. Error bars, shown as the filled area, are obtained from the standard deviation of five runs with different training datasets.

In Fig. 10, the calculated number of MTJ structures required to obtain the best solution is plotted as a function of coefficient of penalty term,  $\alpha$ , of Eq. (4). The results of FM + QA and FM + SA are compared. In the case of  $\Delta E_{\text{total}}$  and TMR [Figs. 10(a) and 10(b)], the number of calculated structures of FM + QA hardly changes against

the  $\alpha$ , but this is not the case in the FM + SA, and hence the searching efficiency of the FM + QA is superior to the FM + SA regardless of the  $\alpha$  value. On the other hand, in the results of RA, a rather large change in the number of calculated structures is obtained and the FM + QA is not superior to the FM + SA [see blue and orange in

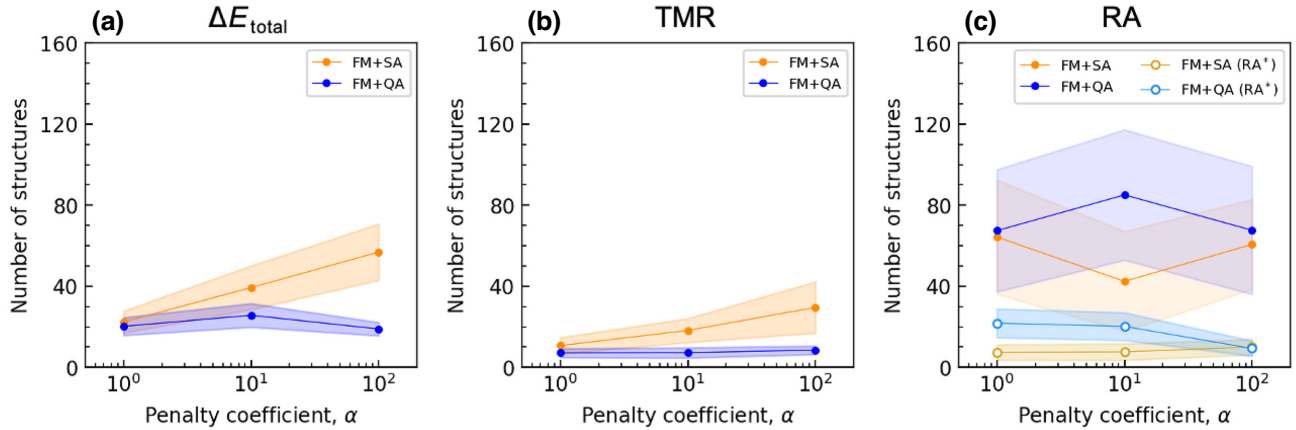


FIG. 10. Calculated number of structures required to obtain the best MTJ structure of (a)  $\Delta E_{\text{total}}$ , (b) TMR, and (c) RA (log scale) as a function of penalty coefficient,  $\alpha$ , where results of FM + QA (blue) and FM + SA (orange) are compared. In (c), results of RA\* are also plotted by opened circles. Error bars, shown as filled area, are obtained from standard deviation of ten runs with different training datasets.

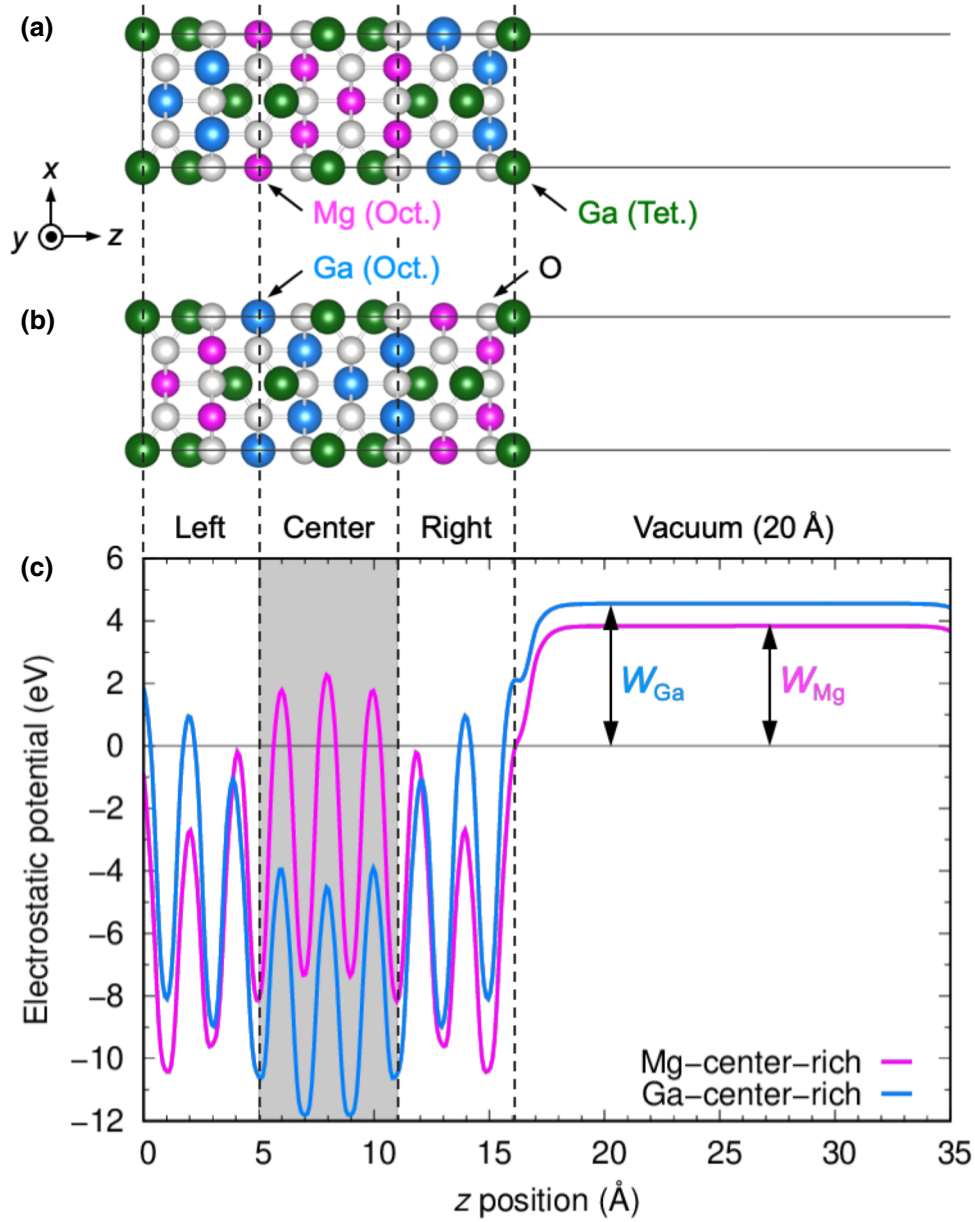


FIG. 11. Slab structures of (a) Mg-center-rich and (b) Ga-center-rich distributions where vacuum region of 20 Å is attached. (c) Layer-averaged electrostatic potential along  $z$  position for Mg-center-rich (magenta) and Ga-center-rich (sky-blue) distributions with corresponding work functions,  $W_X$  ( $X = \text{Mg}$  and  $\text{Ga}$ ). The zero energy of electrostatic potential is set to Fermi energy of each slab.

Fig. 10(c)]. If the searching process is stopped once one of the top-1–top-6 structures is obtained, namely  $\text{RA}^*$ , the dependence of number of calculated structures becomes almost monotonic with respect to the  $\alpha$  [see sky blue and dark orange in Fig. 10(c)]. In this case, the searching efficiency of FM + QA is comparable to that of FM + SA for larger  $\alpha$  value ( $\alpha = 100$ ), and may eventually become better. This result indicates the relevance of choice of  $\alpha$  for efficiently solving the Ising Hamiltonian, however, in the main text, the FM + QA and FM + SA are demonstrated by fixing as  $\alpha = 1$  for simplicity.

## APPENDIX B: WORK-FUNCTION CALCULATIONS

To evaluate the work functions, two slab models consisting of 17 ML of inverse-spinel MGO including the vacuum region of 20 Å are prepared; one is the Mg-center-rich distribution and the other is the Ga-center-rich distribution [see Figs. 11(a) and 11(b)]. Note that we use the 17-ML MGO layers that corresponds to twice that of the unit cell of the spinel structure, for seeing a clear difference of the work functions in these models, and for the vacuum region of 20 Å is confirmed

to be enough to obtain the vacuum energy level precisely for the MGO composed by light elements. The electronic structures are obtained by first-principles calculations using the projected-augmented-wave (PAW) pseudopotentials [52,53] with the GGA [39], implemented in the Vienna *ab-initio* simulation package (VASP) [54]. The atomic positions are fully relaxed using  $11 \times 11 \times 1$   $k$ -point mesh. In Fig. 11(c), layer-averaged electrostatic potential along  $z$  direction is compared between the Mg-center-rich and Ga-center-rich cation distributions. The work functions, defined by a difference between Fermi energy (zero energy) and the electrostatic potential at the vacuum region, result in  $W_{\text{Ga}} = 4.6$  eV for the Ga-center-rich distribution and  $W_{\text{Mg}} = 3.8$  eV for the Mg-center-rich distribution, respectively.

- 
- [1] D. Evans, The Internet of Things: How the Next Evolution of the Internet Is Changing Everything, Cisco White Paper (Cisco Systems, Inc., April 2011), [http://www.cisco.com/web/about/ac79/docs/innov/IoT\\_IBSG\\_0411FINAL.pdf](http://www.cisco.com/web/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf).
- [2] S. A. Wolf, J. Lu, M. R. Stan, E. Chen, and D. M. Treger, The promise of nanomagnetism and spintronics for future logic and universal memory, *Proc. IEEE* **98**, 2155 (2010).
- [3] T. Hanyu, T. Endoh, D. Suzuki, H. Koike, Y. Ma, N. Onizawa, M. Natsui, S. Ikeda, and H. Ohno, Standby-power-free integrated circuits using MTJ-based VLSI computing, *Proc. IEEE* **104**, 1844 (2016).
- [4] Jian-Gang Zhu, Magnetoresistive random access memory: The path to competitiveness and scalability, *Proc. IEEE* **96**, 1786 (2008).
- [5] T. Endoh, H. Koike, S. Ikeda, T. Hanyu, and H. Ohno, An overview of nonvolatile emerging memories—spintronics for working memories, *IEEE J. Emerg. Sel. Top. Circuits Syst.* **6**, 109 (2016).
- [6] S. Bhatti, R. Sbiaa, A. Hirohata, H. Ohno, S. Fukami, and S. N. Piramanayagam, Spintronics based random access memory: A review, *Mater. Today* **20**, 530 (2017).
- [7] T. Miyazaki and N. Tezuka, Giant magnetic tunneling effect in Fe/Al<sub>2</sub>O<sub>3</sub>/Fe junction, *J. Magn. Magn. Mater.* **139**, L231 (1995).
- [8] M. I. Katsnelson, V. Yu. Irkhin, L. Chioncel, A. I. Lichtenstein, and R. A. de Groot, Half-metallic ferromagnets: from band structure to many-body effects, *Rev. Mod. Phys.* **80**, 315 (2008).
- [9] C. Felser and A. Hirohata, editors, *Heusler Alloys: Properties, Growth, Applications* (Springer International Publishing, Cham, 2016), Vol. 222.
- [10] S. S. P. Parkin, C. Kaiser, A. Panchula, P. M. Rice, B. Hughes, M. Samant, and S.-H. Yang, Giant tunnelling magnetoresistance at room temperature with MgO (100) tunnel barriers, *Nat. Mater.* **3**, 862 (2004).
- [11] S. Yuasa, T. Nagahama, A. Fukushima, Y. Suzuki, and K. Ando, Giant room-temperature magnetoresistance in single-crystal Fe/MgO/Fe magnetic tunnel junctions, *Nat. Mater.* **3**, 868 (2004).
- [12] S. Ikeda, J. Hayakawa, Y. Ashizawa, Y. M. Lee, K. Miura, H. Hasegawa, M. Tsunoda, F. Matsukura, and H. Ohno, Tunnel magnetoresistance of 604% at 300 K by suppression of Ta diffusion in CoFeB/MgO/CoFeB pseudo-spin-valves annealed at high temperature, *Appl. Phys. Lett.* **93**, 082508 (2008).
- [13] S. Ikeda, K. Miura, H. Yamamoto, K. Mizunuma, H. D. Gan, M. Endo, S. Kanai, J. Hayakawa, F. Matsukura, and H. Ohno, A perpendicular-anisotropy CoFeB–MgO magnetic tunnel junction, *Nat. Mater.* **9**, 721 (2010).
- [14] 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).
- [15] J. Mathon and A. Umerski, Theory of tunneling magnetoresistance of an epitaxial Fe/MgO/Fe(001) junction, *Phys. Rev. B* **63**, 220403 (2001).
- [16] D. L. Li, Q. L. Ma, S. G. Wang, R. C. C. Ward, T. Hesjedal, X.-G. Zhang, A. Kohn, E. Amsellem, G. Yang, J. L. Liu, J. Jiang, H. X. Wei, and X. F. Han, Controlling spin-dependent tunneling by bandgap tuning in epitaxial rocksalt MgZnO films, *Sci. Rep.* **4**, 7277 (2015).
- [17] Ikhtiar, S. Kasai, P.-H. Cheng, T. Ohkubo, Y. K. Takahashi, T. Furubayashi, and K. Hono, Magnetic tunnel junctions with a rock-salt-type Mg<sub>1-x</sub>Ti<sub>x</sub>O barrier for low resistance area product, *Appl. Phys. Lett.* **108**, 242416 (2016).
- [18] R. Shan, H. Sukegawa, W. H. Wang, M. Kodzuka, T. Furubayashi, T. Ohkubo, S. Mitani, K. Inomata, and K. Hono, Demonstration of Half-Metallicity in Fermi-Level-Tuned Heusler Alloy Co<sub>2</sub>FeAl<sub>0.5</sub>Si<sub>0.5</sub> at Room Temperature, *Phys. Rev. Lett.* **102**, 246601 (2009).
- [19] H. Sukegawa, H. Xiu, T. Ohkubo, T. Furubayashi, T. Niizeki, W. Wang, S. Kasai, S. Mitani, K. Inomata, and K. Hono, Tunnel magnetoresistance with improved bias voltage dependence in lattice-matched Fe/spinel MgAl<sub>2</sub>O<sub>4</sub>/Fe(001) junctions, *Appl. Phys. Lett.* **96**, 212505 (2010).
- [20] H. Sukegawa, Y. Miura, S. Muramoto, S. Mitani, T. Niizeki, T. Ohkubo, K. Abe, M. Shirai, K. Inomata, and K. Hono, Enhanced tunnel magnetoresistance in a spinel oxide barrier with cation-site disorder, *Phys. Rev. B* **86**, 184401 (2012).
- [21] Y. Miura, S. Muramoto, K. Abe, and M. Shirai, First-principles study of tunneling magnetoresistance in Fe/MgAl<sub>2</sub>O<sub>4</sub>/Fe(001) magnetic tunnel junctions, *Phys. Rev. B* **86**, 024426 (2012).
- [22] H. Sukegawa, Y. Kato, M. Belmoubarik, P.-H. Cheng, T. Daibou, N. Shimomura, Y. Kamiguchi, J. Ito, H. Yoda, T. Ohkubo, S. Mitani, and K. Hono, MgGa<sub>2</sub>O<sub>4</sub> Spinel barrier for magnetic tunnel junctions: Coherent tunneling and low barrier height, *Appl. Phys. Lett.* **110**, 122404 (2017).
- [23] K. Masuda and Y. Miura, Bias voltage effects on tunneling magnetoresistance in Fe/MgAl<sub>2</sub>O<sub>4</sub>/Fe(001) junctions: Comparative study with Fe/MgO/Fe(001) junctions, *Phys. Rev. B* **96**, 054428 (2017).
- [24] K. Nawa, K. Masuda, and Y. Miura, Interfacial resonant tunneling induced by folded bands and providing highly spin-polarized current in spinel-oxide barrier junctions, *Phys. Rev. B* **102**, 144423 (2020).
- [25] K. Masuda and Y. Miura, First-principles study on magnetic tunneling junctions with semiconducting CuInSe<sub>2</sub>

- and CuGaSe<sub>2</sub> barriers, *Jpn. J. Appl. Phys.* **56**, 020306 (2017).
- [26] T. Ueno, T. D. Rhone, Z. Hou, T. Mizoguchi, and K. Tsuda, COMBO: An efficient Bayesian optimization library for materials science, *Mater. Discovery* **4**, 18 (2016).
- [27] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. de Freitas, Taking the human out of the loop: A review of Bayesian optimization, *Proc. IEEE* **104**, 148 (2016).
- [28] T. M. Dieb, S. Ju, K. Yoshizoe, Z. Hou, J. Shiomi, and K. Tsuda, MDTs: Automatic complex materials design using Monte Carlo tree search, *Sci. Technol. Adv. Mater.* **18**, 498 (2017).
- [29] W. Paszkowicz, Genetic algorithms, a nature-inspired tool: Survey of applications in materials science and related fields, *Mater. Manuf. Process.* **24**, 174 (2009).
- [30] M. W. Johnson, *et al.*, Quantum annealing with manufactured spins, *Nature* **473**, 194 (2011).
- [31] T. Kadowaki and H. Nishimori, Quantum annealing in the transverse Ising model, *Phys. Rev. E* **58**, 5355 (1998).
- [32] S. Tanaka, R. Tamura, and B. K. Chakrabarti, *Quantum Spin Glasses, Annealing and Computation* (Cambridge University Press, Cambridge, 2017).
- [33] S. Boixo, T. Albash, F. M. Spedalieri, N. Chancellor, and D. A. Lidar, Experimental signature of programmable quantum annealing, *Nat. Commun.* **4**, 2067 (2013).
- [34] K. Kitai, J. Guo, S. Ju, S. Tanaka, K. Tsuda, J. Shiomi, and R. Tamura, Designing metamaterials with quantum annealing and factorization machines, *Phys. Rev. Res.* **2**, 013319 (2020).
- [35] K. Hatakeyama-Sato, T. Kashikawa, K. Kimura, and K. Oyaizu, Tackling the challenge of a huge materials science search space with quantum-inspired annealing, *Adv. Intell. Syst.* **3**, 2000209 (2021).
- [36] S. Ju, Y. Miura, K. Yamamoto, K. Masuda, K. Uchida, and J. Shiomi, Machine learning analysis of tunnel magnetoresistance of magnetic tunnel junctions with disordered MgAl<sub>2</sub>O<sub>4</sub>, *Phys. Rev. Res.* **2**, 023187 (2020).
- [37] H. Sukegawa, S. Mitani, T. Ohkubo, K. Inomata, and K. Hono, Low-resistive monocrystalline Mg-Al-O barrier magnetic tunnel junctions for spin-transfer magnetization switching, *Appl. Phys. Lett.* **103**, 142409 (2013).
- [38] S. Smidstrup, *et al.*, QuantumATK: An integrated platform of electronic and atomic-scale modelling tools, *J. Phys. Condens. Matter* **32**, 015901 (2020).
- [39] J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized Gradient Approximation Made Simple, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [40] K. Nawa, K. Masuda, and Y. Miura, Enhanced Magnetoresistance under Bias Voltage in Fe/MgO/MgAl<sub>2</sub>O<sub>4</sub>/MgO/Fe Trilayer Tunneling Barrier Junction, *Phys. Rev. Appl.* **16**, 044037 (2021).
- [41] K. Masuda, H. Itoh, Y. Sonobe, H. Sukegawa, S. Mitani, and Y. Miura, Interfacial giant tunnel magnetoresistance and bulk-induced large perpendicular magnetic anisotropy in (111)-oriented junctions with fcc ferromagnetic alloys: A first-principles study, *Phys. Rev. B* **103**, 064427 (2021).
- [42] The D-Wave quantum annealing is performed by using software development kits (SDKs) of the Ocean software (D-Wave's Ocean SDK) [available on GitHub at <https://github.com/dwavesystems/dwave-ocean-sdk>, accessed January 1, 2023] and the Amazon Braket Ocean Plugin [available on GitHub at <https://github.com/aws/amazon-braket-ocean-plugin-python>, accessed 1 January 2023].
- [43] FMQA is available on GitHub at <https://github.com/tsudalab/fmqa/> (accessed 1 January 2023).
- [44] M. Zaman, K. Tanahashi, and S. Tanaka, PyQUBO: Python library for mapping combinatorial optimization problems to QUBO form, *IEEE Trans. Comput.* **71**, 838 (2022).
- [45] K. Tanahashi, S. Takayanagi, T. Motohashi, and S. Tanaka, Application of Ising machines and a software development for Ising machines, *J. Phys. Soc. Jpn.* **88**, 061010 (2019).
- [46] For all QA calculations, annealing\_time, the duration of annealing to obtain one QA solution, is set to 2  $\mu$ s, and num\_reads, the number of output solutions in one QA cycle, is set to 200. This means that the calculation to obtain a QA solution annealed over 2  $\mu$ s is repeated 200 times, and the final QA solution is determined as the solution with the lowest Ising energy out of 200 annealing calculations in each QA cycle of Fig. 1(b). For all the other hyperparameters, we used default values in Python libraries of FMQA and D-Wave's Ocean SDK, the latter of which is integrated in PyQUBO. Details of implementation of Python scripts for FM and QA are available in Refs. [44,45].
- [47] The SA and the ES are implemented by the Ocean software [available on GitHub at <https://github.com/dwavesystems/dwave-ocean-sdk>, accessed 1st January 2023] and the BO is by the COMBO3 libraries [26].
- [48] V. S. Denchev, S. Boixo, S. V. Isakov, N. Ding, R. Babush, V. Smelyanskiy, J. Martinis, and H. Neven, What Is the Computational Value of Finite-Range Tunneling?, *Phys. Rev. X* **6**, 031015 (2016).
- [49] T. Graß and M. Lewenstein, Hybrid annealing: Coupling a quantum simulator to a classical computer, *Phys. Rev. A* **95**, 052309 (2017).
- [50] N. Chancellor, Modernizing quantum annealing using local searches, *New J. Phys.* **19**, 023024 (2017).
- [51] M. H. S. Amin, D. V. Averin, and J. A. Nesteroff, Decoherence in adiabatic quantum computation, *Phys. Rev. A* **79**, 022107 (2009).
- [52] P. E. Blöchl, Projector augmented-wave method, *Phys. Rev. B* **50**, 17953 (1994).
- [53] G. Kresse and D. Joubert, From ultrasoft pseudopotentials to the projector augmented-wave method, *Phys. Rev. B* **59**, 1758 (1999).
- [54] G. Kresse and J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, *Phys. Rev. B* **54**, 11169 (1996).