Version of Record: `https://doi.org/10.1016/j.commatsci.2026.114718`
Computational Materials Science, Volume 269, 20 May 2026, 114718. Manuscript Number: COMMAT-D-26-00133
DOI: 10.1016/j.commatsci.2026.114718; Full length article

Analysis of Magnetocrystalline anisotropy from atom pair statistics in tetragonally strained Fe-based alloys: A first-principles and machine learning study

Wen-Chuang Shang[1] and Yoshihiro Gohda[1, *]
[1] *Department of Materials Science and Engineering, Institute of Science Tokyo, Yokohama 226-8501, Japan*

We present a compact physics-guided machine-learning framework for predicting and interpreting the magnetocrystalline anisotropy (MCA) energy of tetragonally strained Fe-based alloys. Using density functional theory (DFT) with spin–orbit coupling (SOC), we generated $\sim 10^3$ strained structures spanning broad ranges of volume V and tetragonality c/a. The core model is a ridge regressor built on symmetry-aware atom-pair statistics, where species-resolved pair distances, orientations, and counts are organized into channels. Trained on 2×2×2 Fe–Si supercells at four Si concentrations, the model reproduces the main DFT trends and is further assessed on a held-out 3×3×3 $\mathrm{Fe_{0.889}Si_{0.111}}$ supercell as a direct size/composition generalization check. The learned anisotropy coefficient functions quantify channel-resolved contributions as a function of pair distance, enabling distance- and orientation-resolved interpretation of strain-dependent MCA within a pair-additive surrogate. A random-forest baseline using only global descriptors $(V, c/a)$ highlights the advantage of local pair statistics for resolving chemical and geometrical effects. Extension to ordered Fe–X (X = Ge, Ga, Al) illustrates how equilibrium-volume changes correlate with MCA trends, with the averaged exchange splitting serving as a useful intermediate quantity.

Keywords: Magnetoelastic effect; Spin–orbit coupling; Density functional theory; Machine learning

1. Introduction

Magnetostriction — the spontaneous change in shape or length of a magnetic material upon rotation of its magnetization — is a direct consequence of spin–lattice coupling mediated by the spin–orbit coupling (SOC) [1, 2]. Its magnitude, measured as the strain, spans a striking range of 10^{-8}–10^{-3}, depending on chemistry and symmetry [3]. A large response is desired for sensors, actuators, energy harvesters, and spintronic devices [3–7], whereas transformer cores (electric steels) require negligible magnetostriction to suppress acoustic noise and power loss, which is achieved near 6.5 mass% (∼ 13 at.%) Si in Fe–Si alloys, where theoretical and experimental studies consistently report negligible magnetostriction [8–11]. Magnetostriction depends on both elastic properties and magnetoelastic coupling [2, 12]. Because magnetoelastic coupling is directly governed by the SOC, it is more complex and is intimately linked to the strain dependence of the magnetocrystalline anisotropy (MCA) energy E_{MCA}—the energy difference between different magnetization directions. Understanding how E_{MCA} responds to strain is essential for designing magnetostrictive materials.

For cubic crystal systems, a prototypical strain mode relevant to magnetostriction is the tetragonal strain. To evaluate the magnetostriction along the ⟨001⟩ direction at saturation magnetization, the coefficient $\lambda_{001} \equiv \Delta z/z$ is commonly employed. This is typically implemented under a constant-volume condition: applying fractional strain ε_z along the z axis while preserving the total volume V. The coefficient λ_{001} is then obtained from the ε_z dependence of both the total energy E_{total} and E_{MCA} [2, 5, 13]:

$$\lambda_{001} = \frac{2\,\mathrm{d}E_{\mathrm{MCA}}/\mathrm{d}\varepsilon_z}{3\,\mathrm{d}^2 E_{\mathrm{total}}/\mathrm{d}\varepsilon_z^2} = -\frac{b_{\mathrm{me}}}{3C'}\,, \tag{1}$$

where $E_{\mathrm{MCA}} = E([100]) - E([001])$ (a convention often used for tetragonal magnetostriction); the numerator and denominator correspond, respectively, to the magnetoelastic coupling coefficient $b_{\mathrm{me}} = (-2/3)\,V^{-1}\,\mathrm{d}E_{\mathrm{MCA}}/\mathrm{d}\varepsilon_z$ and the tetragonal shear modulus $C' = (1/3)\,V^{-1}\,\mathrm{d}^2E_{\mathrm{total}}/\mathrm{d}\varepsilon_z^2$. This study focuses on the numerator term of (1), which captures the strain-induced variation in E_{MCA}.

In practice, within density functional theory (DFT), E_{MCA} can be evaluated either (i) by treating the SOC term as a perturbation and evaluating the second-order energy correction, or (ii) by adding the SOC term directly to the Kohn–Sham equations and performing DFT calculations for different magnetization orientations. The perturbative expression naturally decomposes E_{MCA} into four spin-resolved channels: two spin-conserving (↑↑, ↓↓) and two spin-flip (↑↓, ↓↑). Within the second-order SOC perturbative framework, the ↓↓ contribution corresponds

---

* gohda@mct.isct.ac.jp

to the Bruno term, which is linked to orbital-moment anisotropy [14], whereas, under certain approximations, the $\uparrow\downarrow$ channel yields the van der Laan term, associated with the spin-density quadrupole [15, 16] (details in Appendix A).

In several representative Fe-based binary alloys, first-principles studies have revealed several prototypical behaviors. In $Fe_{100-x}Ga_x$ (Galfenol), λ_{001} grows rapidly up to $x \sim 17$–$20\,\%$, drops sharply near $x \sim 19$–$25\,\%$, and can be resurrected by suitable short/medium-range ordering or targeted ternary doping [4, 5, 17–19]. In $Fe_{1-x}Ge_x$, λ_{001} turns negative beyond $x \sim 14\,\%$ as $D0_3$ order develops and e_g holes emerge above the Fermi level [18]. In $Fe_{1-x}Si_x$, experimental and theoretical studies indicate that λ_{001} follows a broadly decreasing trajectory as x increases, tending to vanish around $x \approx 0.10$–0.13, though local non-monotonic features with small upturns are observed near certain concentrations [8–11]. In particular, for $Fe_{0.889}Si_{0.111}$, a DFT study reported that the exceptional near-zero magnetostriction arises from the mutual cancellation of sizable contributions from multiple types of neighboring atoms extending up to the 10th nearest neighbors, none of which individually dominate the overall response [8]. In parallel, first-principles studies on platforms beyond bulk Fe-based binaries have broadened the landscape of anisotropy engineering. These works have elucidated strain-driven anisotropy in Heusler systems [20], domain-controlled anisotropy in multiferroic heterostructures where ferroelectric patterns modulate magnetic anisotropy [21], and interfacial magnetoelectric coupling at $Fe_3Si/BaTiO_3$ (001) and $Fe/BaTiO_3$ (001) interfaces [22, 23]. Related efforts have also addressed strain effects on anisotropy in rare-earth magnets [24], while more broadly, surface and interface magnetism relevant to anisotropy control has been explored [25].

Most DFT studies investigate a few hand-picked phase prototypes for each alloy separately and analyze the E_{MCA}–strain relation on a case-by-case basis. This strategy provides useful interpolation within a fixed system, but does not always reliably yield generally transferable rules across compositions, which limits its utility for prospective material design. Thus, beyond conventional DFT calculations, one solution has been to employ machine learning (ML) models to learn shared behaviors across material families.

Early frameworks such as the Behler–Parrinello neural network demonstrated that descriptor-based models fitted to DFT data can reproduce interatomic forces and energies with near-first-principles accuracy [26]. More recently, a DFT-informed artificial neural network (ANN) potential for body-centered cubic (bcc) iron has shown excellent accuracy in property predictions while retaining high computational efficiency [27]. In parallel, graph neural networks (GNNs) have become prominent by learning structure–property relationships directly from atomic graphs, with nodes representing atoms and edges encoding interatomic distances. Examples include SchNet and the Crystal Graph Convolutional Neural Network (CGCNN), which have successfully predicted various material properties using DFT databases [28, 29]. Building on these ideas, the Atomistic Line Graph Neural Network (ALIGNN) further incorporates bond-angle correlations through a line-graph architecture, highlighting the importance of angular information to improve model accuracy [30]. However, few of these ML models have yet been applied to predict the MCA energy, mainly because existing databases contain very limited DFT data with the SOC.

More recently, several DFT–ML studies have incorporated the SOC to enable direct predictions of E_{MCA} in selected systems [31, 32]. These efforts employed multiple data samples across different compositions or configurations; however, each composition or configuration was typically represented by a single equilibrium geometry, without systematically exploring how the MCA responds under strain. Recent perspective analyses have emphasized that intrinsic magnetic performance is governed by electronic-structure descriptors such as exchange splitting, orbital anisotropy near the Fermi level, and spin–orbit-coupling-active states, while descriptor design remains a key bottleneck in machine-learning-guided magnet discovery [33]. More broadly, what remains limited is a compact and interpretable framework that connects atomic composition/geometry to strain-dependent MCA, while providing a transparent basis for assessing generalization beyond the systems used for training.

In this work, we propose a compact DFT–ML framework centered on symmetry-aware atom-pair statistics to analyze the strain dependence of E_{MCA} in bulk Fe-based alloys. The pair-based formulation provides length- and orientation-resolved attributions of magnetoelastic coupling. We emphasize that the pair-additive model is used as an interpretable surrogate: the pair-based descriptor and the associated framework are formally extensible and therefore have potential applicability to other systems, although predictive accuracy outside the present validation scope must still be established case by case.

This paper is organized as follows. Section 2 describes the DFT workflow, strain protocol, the construction of local and global descriptors, and the models used in this work. Section 3 presents the main numerical results and key discussions. Section 4 concludes with the main findings.

2. Computational details

2.1. DFT calculation

Eight Fe-rich ordered bcc-derived systems were considered. For Fe–Si alloys, we sampled 6.25, 12.5, 18.75, and 25 at.% Si with 2×2×2 supercells, and 11.1 at.% Si with a 3×3×3 supercell [Fig. 1]. In addition, three compounds at 25 at.% — $Fe_{0.75}Ge_{0.25}$, $Fe_{0.75}Al_{0.25}$, and $Fe_{0.75}Ga_{0.25}$ — were included. For each of the seven $2 \times 2 \times 2$ compositions, approximately 200 strained states were gener-

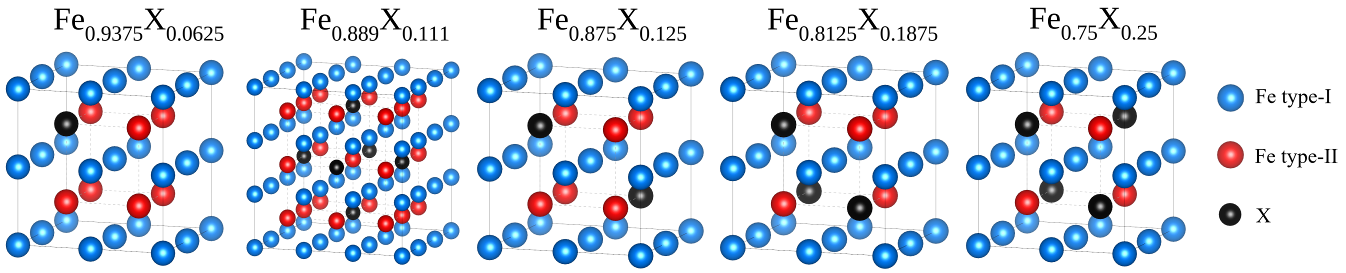


Fig. 1. Supercells derived from the bcc lattice employed in this study. The Si series includes $Fe_{0.9375}Si_{0.0625}$, $Fe_{0.889}Si_{0.111}$, $Fe_{0.875}Si_{0.125}$, $Fe_{0.8125}Si_{0.1875}$, and $Fe_{0.75}Si_{0.25}$ ($D0_3$–Fe_3Si). For comparison across congeners, three additional $D0_3$ compounds $Fe_{0.75}Ge_{0.25}$, $Fe_{0.75}Al_{0.25}$, and $Fe_{0.75}Ga_{0.25}$ were also investigated.

ated by varying the cell volume V and the tetragonal distortion represented by c/a. Specifically, V was randomly sampled in the range of [18, 28] Å^3 (per 1×1×1 bcc unit), while c/a was randomly varied in the range [0.92, 1.08]. For the larger 3×3×3 $Fe_{0.889}Si_{0.111}$ cell, 50 strained states were sampled within the same (V, c/a) range.

The 25 at.% compositions we consider correspond to the $D0_3$ compound, a stable ordered phase reported in multiple Fe-based binary systems [34–40]. Intermediate $Fe_{1-x}Si_x$ compositions (6.25–18.75 at.%) were modeled by ordered supercells rather than random A2 solid solutions; these configurations preserve the cubic symmetry and are stable under tetragonal strain. This choice is practically feasible and aims to simulate mixed phases, reflecting both computational tractability and physical realism: The different ordered phases A2/B2/$D0_3$ may coexist rather than perfect randomness [41–44], and similar ordered configuration strategies have been adopted in DFT studies of Fe–Si, Fe–Ge, Fe–Be to capture magnetoelastic trends [8, 18, 45].

DFT calculations were performed using the Vienna Ab initio Simulation Package (VASP) [46, 47], which implements the projector augmented-wave (PAW) method [48]. Exchange–correlation effects were treated within the generalized gradient approximation (GGA) using the Perdew–Burke–Ernzerhof (PBE) functional [49–51]. The plane-wave cutoff energy was set to 350 eV. Brillouin-zone integrations used Monkhorst–Pack meshes [52] of 15 × 15 × 15 for 2 × 2 × 2 bcc cells and 9 × 9 × 9 for 3 × 3 × 3 bcc cells. The SOC was included in the evaluation of E_{MCA}. The self-consistent field (SCF) tolerance was 10^{-8} eV. To isolate the imposed (V, c/a) dependence, all calculations were performed in a one-shot manner at the prescribed (V, c/a) conditions without relaxations. Additional fixed-cell internal-relaxation checks indicate no changes for the $\mathrm{Fe}_{0.875}X_{0.125}$ and $\mathrm{Fe}_{0.75}X_{0.25}$ compositions examined here as their atomic arrangements are more strongly constrained by symmetry, and only slight changes for $Fe_{0.9375}Si_{0.0625}$, $Fe_{0.8125}Si_{0.1875}$ and $\mathrm{Fe}_{0.889}X_{0.111}$, with the internal-coordinate changes showing no pronounced dependence on tetragonal strain and with shifts in E_{MCA} that fall within the DFT numerical-accuracy range and are therefore not consequential for assessing the overall strain-dependent trend.

As noted earlier, two alternative approaches are commonly employed to evaluate E_{MCA} [53, 54]. (i) The first approach is based on second-order perturbation theory, where the SOC is treated as a perturbation on the unperturbed eigenstates obtained from non-SOC calculation (with spin aligned along a quantization axis). The second-order correction $E^{(2)}_{\mathrm{SOC}}$ to the energy is evaluated for magnetization along [001] and [100], and their difference gives $E^{(2)}_{\mathrm{MCA}}$. This perturbative formulation further enables decomposition into spin-conserving ($\uparrow\uparrow, \downarrow\downarrow$) and spin-flip ($\uparrow\downarrow, \downarrow\uparrow$) contributions [14–16]. Implementation details are provided in Appendix A; (ii) The second approach is a direct SOC-included DFT calculation, where spin–orbit coupling is explicitly incorporated into the Kohn–Sham Hamiltonian [53–55]. The total energies $E([001])$ and $E([100])$ are then evaluated from the resulting Kohn–Sham wavefunctions. In practice, E_{MCA} can be computed using a frozen-density protocol, in which a self-consistent non-SOC calculation is first performed to converge the charge density, and non-self-consistent SOC calculations for different magnetization orientations are then carried out using the fixed density and potential; the magnetocrystalline anisotropy energy is finally evaluated as $E_{\mathrm{MCA}} = E([100]) - E([001])$. In the present work, E_{MCA} is evaluated from direct total-energy differences within DFT including the SOC (approach (ii)), while the corresponding perturbative decomposition and the definitions of $\mathcal{V}$ and $\mathcal{B}$ are given in Appendix A and used as additional learning targets.

2.2. The statistical descriptors for local features of atom pairs: Physics-guided pair-based ridge regression

We focus on bulk Fe–Si alloys to construct an interpretable surrogate for predicting an anisotropic target Y_{ani} from geometric descriptors of atomic structures, characterized by the counts, lengths, and orientations of atom pairs. The training set consists of

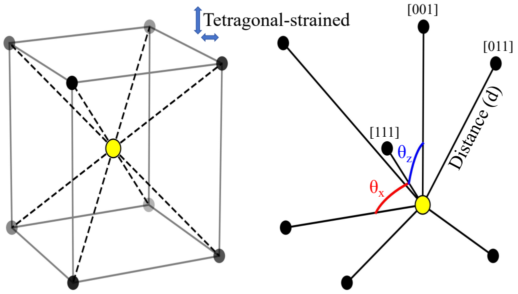


Fig. 2. Tetragonally strained bcc structure and the atom pairs with different orientations.

Table 1. Orientation classes and thresholds used for classification of tetragonal systems. The rightmost column lists the total number of unique geometric pairs per class in a generic $n \times n \times n$ bcc cell.

Class	Angular condition	d cutoff (Å)	Count
111	all $\theta_{x,y,z}$ satisfy $\lvert\theta - 54.7°\rvert < 4°$	< 3.0	$8n^3$
100	$\lvert\theta_x\rvert < 4°$, $\lvert\theta_z - 90°\rvert < 4°$	< 3.5	$2n^3$
001	$\lvert\theta_z\rvert < 4°$, $\lvert\theta_x - 90°\rvert < 4°$	< 3.5	$2n^3$
010	$\lvert\theta_y\rvert < 4°$, $\lvert\theta_{x,z} - 90°\rvert < 4°$	< 3.5	$2n^3$
110	$\lvert\theta_x - 45°\rvert < 4°$, $\lvert\theta_y - 45°\rvert < 4°$	< 4.5	$4n^3$
101	$\lvert\theta_x - 45°\rvert < 4°$, $\lvert\theta_z - 45°\rvert < 4°$	< 4.5	$4n^3$
011	$\lvert\theta_y - 45°\rvert < 4°$, $\lvert\theta_z - 45°\rvert < 4°$	< 4.5	$4n^3$

Table 2. Counts of accepted atom pairs per cell in the four merged orientation groups, obtained from the mode of the counts in the dataset (PBC-canonicalized). Entries are "Fe–Fe/Fe–Si". Si–Si pairs are not included in this table.

Composition	111	100/001/010	110/101/011
$Fe_{0.9375}Si_{0.0625}$ $(2 \times 2 \times 2)$	56/8	42/6	84/12
$Fe_{0.875}Si_{0.125}$ $(2 \times 2 \times 2)$	48/16	36/12	72/24
$Fe_{0.8125}Si_{0.1875}$ $(2 \times 2 \times 2)$	40/24	30/18	72/12
$Fe_{0.75}Si_{0.25}$ $(2 \times 2 \times 2)$	32/32	24/24	72/0
$Fe_{0.889}Si_{0.111}$ $(3 \times 3 \times 3)$	168/48	129/30	264/48

DFT data from 2×2×2 bcc supercells at four concentrations $Fe_{0.9375}Si_{0.0625}$, $Fe_{0.875}Si_{0.125}$, $Fe_{0.8125}Si_{0.1875}$, and $Fe_{0.75}Si_{0.25}$ [Fig. 1]. To examine generalization across different cell sizes and atomic configurations, we also tested the model on an unseen data set for $Fe_{0.889}Si_{0.111}$, represented by a larger 3×3×3 cell. The ML framework for these atom-pair-based (pair-based) statistical descriptors is summarized as follows.

For each structure, neighboring atoms were identified using a uniform cutoff of 4.5 Å, excluding self-pairs. Each pair was labeled as $(p, q, \mathbf{T})$, where p and q index atoms within the unit cell and $\mathbf{T}$ is a lattice translation vector accounting for periodic boundary conditions (PBC). To avoid double counting, only triplets satisfying $p < q$ were retained. For each valid pair, the bond vector $\mathbf{r}_{pq} = \mathbf{R}_q + \mathbf{T} - \mathbf{R}_p$ was used to compute the bond length $d = |\mathbf{r}_{pq}|$ and its acute angles $(\theta_x, \theta_y, \theta_z)$ with respect to the Cartesian axes. In the bcc structures considered here, accessible pairs include first-nearest neighbors (1NN) along [111] directions, second-nearest neighbors (2NN) along [100]/[010]/[001], and third-nearest neighbors (3NN) along [110]/[101]/[011]. Each unique Fe–Fe or Fe–Si pair was assigned to a class according to the angular and distance thresholds listed in Table 1. We stored $(d, \sin\theta_x, \sin\theta_z)$ for each accepted pair. Note that pairs exactly along [010] do not contribute to Y_{ani} because $\sin^2\theta_x - \sin^2\theta_z = 0$. In addition, Si–Si pairs are excluded from the descriptor set because Si carries negligible local moment in these Fe-rich alloys and is therefore expected to contribute weakly to the anisotropy targets considered here (see Fig. 2).

By construction, an "orientation class" aggregates directions that differ only by sign inversions of their Miller indices (e.g., [101] and [10$\bar{1}$]), while keeping the index order fixed. This reflects the physical equivalence under tetragonal symmetry, but does not merge directions related by cyclic index permutations such as [110] and [011], since their $\sin\theta_x$ and $\sin\theta_z$ differ. This convention is consistent with the angular thresholds in Table 1. The small angular tolerances in Table 1 are considered to accommodate slight anisotropic strains and modest atomic displacements without reclassifying pairs, thereby improving robustness. To represent each type of uniaxial pair subject to the same global crystal field but oriented along different local symmetry axes, six pair channels (labeled by t) are defined: $Fe–Fe_{111}$ and $Fe–Si_{111}$ (uniaxial pairs embedded in a C_3-symmetric environment about the pair axis); $Fe–Fe_{100/001}$ and $Fe–Si_{100/001}$ (in a C_4-symmetric environment); $Fe–Fe_{110/101/011}$ and $Fe–Si_{110/101/011}$ (in a C_2-symmetric environment).

Following an anisotropic expansion form [2], the leading contribution of pair i to an anisotropic target Y_{ani} (defined as a difference between magnetization along [100] and [001]) can be expressed as

$$Y_i \approx K_Y^{(t)}(d_i)\left[\sin^2\theta_{x,i} - \sin^2\theta_{z,i}\right], \qquad (2)$$

where t is the index of the pair channel and $K_Y^{(t)}(d)$ is the target-specific anisotropy coefficient for atom pairs in channel t, which absorbs information about pair species and the strain-dependent pair length d. Here, $\theta_{x,i}$ $(\theta_{z,i})$ is the angle between the bond vector of pair i and the Cartesian x (z) axis. The factor $\sin^2\theta_{x,i} - \sin^2\theta_{z,i}$ therefore corresponds to the change in a uniaxial $\sin^2\theta$ anisotropy term when the direction of interest (the magnetization direction in this study) is switched between [100] and [001]. Assuming additive contributions from individual pairs and neglecting cross-pair interactions,

$$Y_{\text{ani}}(n) \approx \sum_{i \in \mathcal{P}(n)} Y_i. \qquad (3)$$

In this study, $Y_{\text{ani}} \in \{E_{\text{MCA}}, \mathcal{V}, \mathcal{B}\}$, where $\mathcal{V}$ and $\mathcal{B}$ denote the van der Laan term and the Bruno term in the second-order SOC decomposition (Appendix A). Let $\mathcal{P}(n)$ denote the set of all accepted interatomic pairs (Fe–Fe or Fe–Si) in structure n that pass the class-dependent angular and distance criteria in Table 1. Each $i \in \mathcal{P}(n)$ is assigned to exactly one channel t. For each channel t, we approximate $K_Y^{(t)}(d)$ using a Gaussian radial-basis function (RBF) expansion

$$K_Y^{(t)}(d) \approx \sum_{\mu \in \mathcal{M}_t} w_\mu^{(t)} \exp\Big[-\tfrac{(d-\mu)^2}{2\sigma^2}\Big], \tag{4}$$

with $\mathcal{M}_t = \{\mu_1^{(t)}, \ldots, \mu_{N_\mu}^{(t)}\}$ the set of N_μ uniformly spaced centers in the channel-specific distance range $[d_{\mathrm{min}}^{(t)}, d_{\mathrm{max}}^{(t)}]$. As (3) describes, summing (2) over pairs in structure n yields a linear model in the weights $w_\mu^{(t)}$ with features

$$X_{n,\mu}^{(t)} = \frac{1}{N_n} \sum_{i \in \mathcal{P}(n) \cap t} \exp\Big[-\tfrac{(d_{n,i}-\mu)^2}{2\sigma^2}\Big] \left[\sin^2\theta_{x,i} - \sin^2\theta_{z,i}\right], \tag{5}$$

and prediction

$$Y_{\mathrm{ani}}^{\mathrm{ML}}(n) = \sum_t \sum_{\mu \in \mathcal{M}_t} w_\mu^{(t)} X_{n,\mu}^{(t)}. \tag{6}$$

Here N_n is the number of atoms in structure n. Dividing by N_n renders the descriptor intensive and makes $Y_{\mathrm{ani}}^{\mathrm{ML}}$ directly comparable to the DFT target averaged per atom (eV/atom), allowing consistent use across supercell sizes without the weights absorbing size factors.

(6) is fitted to $Y_{\mathrm{ani}}^{\mathrm{DFT}}$ using ridge regression [56] by minimizing the regularized least-squares loss

$$\mathcal{L}(\mathbf{w}) = \sum_n \Big(Y_{\mathrm{ani}}^{\mathrm{DFT}}(n) - Y_{\mathrm{ani}}^{\mathrm{ML}}(n)\Big)^2 + \alpha \, \|\mathbf{w}\|_2^2, \tag{7}$$

performed separately for each target E_{MCA}, $\mathcal{V}$ and $\mathcal{B}$, yielding the target-specific anisotropy coefficient functions $K_Y^{(t)}(d)$, where $\mathbf{w}$ collects all coefficients $w_\mu^{(t)}$ across channels and RBF centers. Before regression, all features $X_{n,\mu}^{(t)}$ were standardized by their standard deviation $\mathrm{std}(X_{n,\mu}^{(t)})$ only (without mean subtraction), i.e., $\tilde{X}_{n,\mu}^{(t)} = X_{n,\mu}^{(t)}/\mathrm{std}(X_{n,\mu}^{(t)})$, and the ridge regression model was fitted without an intercept term. By neither centering the features nor fitting an intercept, this approach preserves the physical condition $Y_{\mathrm{ani}} = 0$ for cubic states (for anisotropy differences such as E_{MCA} and its perturbative components) and avoids artificial offsets. We further verified that the descriptor construction and model predictions satisfy the required symmetry relations under [100] ↔ [001] exchange and mirror operations; details are provided in Appendix B.

Unless stated otherwise, we use the baseline hyperparameters $(N_\mu,\ \sigma,\ \alpha) = \big(20,\ 0.02\ \text{Å},\ 30\big)$. To assess hyperparameter sensitivity and to avoid relying on a single-point choice, we performed targeted scans over σ, α, and N_μ, summarized in Appendix C. This baseline setting is selected from a stable region where (i) σ is small enough to resolve strain-induced distance variations but not so small as to over-amplify noise-like fluctuations, (ii) N_μ provides adequate radial resolution, and (iii) the ridge penalty α stabilizes coefficients without severe loss of fidelity. The learned $\{w_\mu^{(t)}\}$ reconstruct channel-wise anisotropy coefficient curves $K_Y^{(t)}(d)$, enabling length-resolved attribution for Fe–Fe versus Fe–Si and for each pair channel.

The present work restricts attention to 1NN–3NN pairs in ordered Fe-rich Fe–Si; to quantify the practical impact of this truncation, we provide additional neighbor-shell ablation results in Appendix D. In formulation, the descriptor construction can be extended to additional chemical species by defining additional symmetry-aware pair channels, and the neighbor range can be increased beyond 3NN by enlarging the accepted pair set and the corresponding distance windows.

2.3. The descriptors for global features (V, c/a): random forest model

For tetragonally strained bcc-derived structures, the global state can be directly parameterized by V and c/a. These two macroscopic variables provide a low-dimensional descriptor set that directly captures the strain dependence of E_{MCA}. In this study, the global analysis is based on seven ordered 2×2×2 bcc supercells representing $\mathrm{Fe_{0.9375}Si_{0.0625}}$, $\mathrm{Fe_{0.875}Si_{0.125}}$, $\mathrm{Fe_{0.8125}Si_{0.1875}}$, $\mathrm{Fe_{0.75}Si_{0.25}}$, $\mathrm{Fe_{0.75}Ge_{0.25}}$, $\mathrm{Fe_{0.75}Al_{0.25}}$, and $\mathrm{Fe_{0.75}Ga_{0.25}}$.

To obtain a nonparametric representation within the span of the training data, we employ a random forest (RF) regressor [57] implemented in *scikit-learn* [58], trained on (V, c/a) together with a nominal composition label comp $\in \{1, \ldots, 7\}$, with all features standardized by z-score transformation prior to fitting. The model uses 300 decision trees, with no restriction on maximum depth; splits require at least two samples, leaves at least one, the number of features considered at each split follows the square-root heuristic, and bootstrap sampling is applied.

It should be emphasized that no independent test set is prepared for these global descriptors; the RF regression model functions purely as an in-sample interpolant to capture systematic trends. Since the composition label merely indexes the seven systems without element-specific information, this surrogate is not intended to generalize to unseen compositions. We use the RF regression as a baseline to benchmark the pair-based ridge regression, which incorporates chemically and geometrically resolved local features for broader applicability, as introduced in Section 2.2.

3. Results and discussion

As a prelude to the detailed results and as a quantity used throughout this study, we define an effective magnetoelastic coupling parameter $E_{\mathrm{MCA}}/(c/a-1)$, a physically motivated measure that approximately captures the sensitivity of the MCA to tetragonal distortion. While the conventional magnetoelastic coupling coefficient b_{me} for tetragonally strained cubic systems is given by $-b_{\mathrm{me}} = (2/3)\,V^{-1}\,\mathrm{d}E_{\mathrm{MCA}}/\mathrm{d}\varepsilon_z$, its interpretation assumes a constant-volume constraint, under which the uniaxial strain ε_z serves as a well-defined measure of tetragonal distortion. However, in our study, the volume V is varied across a broad range, making ε_z inadequate to describe the system's tetragonal geometry. In contrast, the tetragonality $c/a-1$ directly captures the deviation from cubic symmetry independent of volume, making it a more robust quantity for quantifying tetragonal states. The relations between E_{MCA} and $(c/a-1)$ or closely related descriptions and formulas, which quantify the strain derivative of E_{MCA}, have been used — explicitly or implicitly — as a practical proxy for magnetoelastic coupling in various materials systems [20, 59–62]. For numerical stability, we exclude nearly cubic states with $|c/a-1| \leq 0.002$ in this study. Accordingly, besides $E_{\mathrm{MCA}}/(c/a-1)$, we also examine $\mathcal{V}/(c/a-1)$ and $\mathcal{B}/(c/a-1)$ to compare the tetragonally strain responses on the same footing.

We first present the results of the pair-based ridge regression model, which serves as the core framework of this study (Section 3.1). For comparison and complementary insight, the results of the RF regression model based on the global descriptors $(V, c/a)$ are also examined (Section 3.2). Finally, we relate the macroscopic trends $E_{\mathrm{MCA}}/(c/a-1)$–V to the microscopic exchange splitting $\Delta\varepsilon_{\mathrm{ex}}$ to discuss their physical origin (Section 3.3).

3.1. Prediction and analysis for E_{MCA}, $\mathcal{V}$, and $\mathcal{B}$ from the pair-based ridge regression

As introduced in Section 2.2, we trained the pair-based ridge regression model of Eqs. (2)–(6) on 2×2×2 Fe–Si supercells at four concentrations, using only atom pair statistics as inputs, and evaluated its generalization on the unseen 3×3×3 $\mathrm{Fe}_{0.889}\mathrm{Si}_{0.111}$ cell. We trained separate ridge regression models for E_{MCA}, $\mathcal{V}$ and $\mathcal{B}$, using the same input set and same hyperparameters.

Fig. 3(a) shows that the learned map $E^{\mathrm{ML}}_{\mathrm{MCA}}$ reproduces the overall DFT trends for the training set of 2×2×2 structures and yields consistent behavior for the held-out 3×3×3 $\mathrm{Fe}_{0.889}\mathrm{Si}_{0.111}$ test structures. Within the present Fe-rich ordered Fe–Si scope, this indicates that local pair statistics provide a useful and interpretable surrogate for dominant strain-dependent trends, while the moderate test R^2 highlights the limitations of a strict pair-additive approximation. The close agreement between training and test root mean squared error (RMSE) values in Fig. 3 suggests the absence of severe overfitting under the chosen hyperparameters, but it does not imply that all SOC-driven many-body effects are captured. A composition-wise Leave-One-Composition-Out (LOCO) validation is summarized in Appendix E and further clarifies the scope of generalization supported by the present data. Fig. 3(b) and (c) report analogous illustrations for $\mathcal{V}$ and $\mathcal{B}$, respectively. Notably, the model achieves the best predictive performance for $\mathcal{V}$ among the three targets, exhibiting close agreement with the DFT data. This suggests that the local pair descriptors — constructed solely from interatomic distances and orientation angles up to the 3NN shell — capture a substantial part of the strain-dependent anisotropic behavior of the intra-atomic magnetic dipole moment within the adopted surrogate. The presence of less-than-perfect R^2_{test} in Figs. 3(a)–(c) reflects the intrinsic challenge of predicting E_{MCA} for unseen structures and strains, largely due to the simplifying assumptions in (2) and (3), which neglect higher-order pair-pair couplings and environment-dependent effects. In preliminary tests, we extended the model with triplet terms based on $u_x v_x - u_z v_z$, where $\mathbf{u}$ and $\mathbf{v}$ are unit vectors along two neighboring pairs sharing a common anchor atom, and u_x, u_z, v_x, and v_z are their directional cosines with respect to the Cartesian x and z axes, together with an RBF expansion of the distance between the two non-anchor atoms. This triplet augmentation did not yield a clear improvement in held-out performance and reduced the interpretability of the original pair-based model. We therefore retain the present pair-based model as the main interpretable surrogate in this work, while leaving the role of three-body and higher-order terms for future study, potentially through more effective descriptors capable of capturing their subtler contributions.

The learned anisotropy coefficient functions are shown in Figs. 4(a, c, e). For clarity, we use the 1NN-pair channels as representative examples to interpret $K^{(t)}_{\mathrm{Y}}(d)$. For $K^{(t)}_{\mathrm{MCA}}(d)$ [Fig. 4(a)], $K^{(\mathrm{Fe–Fe},111)}_{\mathrm{MCA}}(d)$ remains positive over most of its characteristic distance range, so $\mathrm{Fe–Fe}_{111}$ pairs generally provide a positive contribution to $E_{\mathrm{MCA}}/(c/a-1)$, whereas $\mathrm{Fe–Si}_{111}$ pairs generally yield a negative contribution. For $K^{(t)}_{\mathcal{V}}(d)$ [Fig. 4(c)], both $\mathrm{Fe–Fe}_{111}$ and $\mathrm{Fe–Si}_{111}$ channels are positive over almost the entire pair-length windows, indicating that 1NN pairs contribute positively to $\mathcal{V}/((c/a)-1)$. The trends of $K^{(\mathrm{Fe–Fe},111)}_{\mathcal{B}}(d)$ and $K^{(\mathrm{Fe–Si},111)}_{\mathcal{B}}(d)$ [Fig. 4(e)] are close to those of $K^{(t)}_{\mathrm{MCA}}(d)$, consistent with their similar macroscopic behaviors. Analogous discussions can be extended to the other channels by incorporating the corresponding channel-specific geometry. Figs. 4(b,d,f), constructed from the channel-resolved regression coefficients and averaged over all strained states for each composition, illustrates how different channels combine to determine the strain response of these three targets. In these averaged responses, the contributions from different neighbor shells — 1NN (111), 2NN (100/001), and 3NN (110/101/011) — are of comparable magnitude.

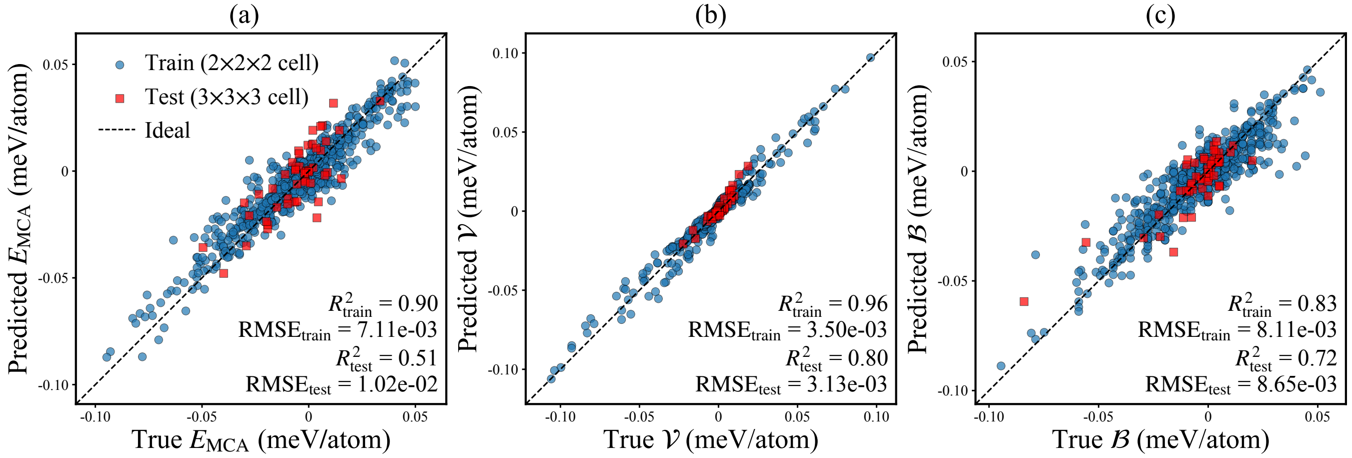


Fig. 3. Out-of-sample evaluation of the pair-based ridge regression model. (a) Predicted vs. true E_{MCA} (meV/atom). (b) Predicted vs. true van der Laan term $\mathcal{V}$. (c) Predicted vs. true Bruno term $\mathcal{B}$. Dots denote training data from 2×2×2 cells and squares denote the unseen 3×3×3 $\mathrm{Fe}_{0.889}\mathrm{Si}_{0.111}$ cell used only for testing. All models are trained with standardized features; the dashed line indicates the ideal reference. Quantitative performance metrics (R^2 and RMSE) are included in the subplots.

An exception occurs for the Fe–Si$_{110/101/011}$ channel in $\mathrm{Fe}_{0.75}\mathrm{Si}_{0.25}$, where the absence of such Fe–Si pairs along these directions leads to a vanishing contribution.

To verify the stability of these regression-based trends, the learned anisotropy coefficients are further assessed within the Bayesian interpretation of ridge regression. Within a Bayesian formulation, we consider the linear model $\boldsymbol{Y}_{\mathrm{ani}}^{\mathrm{DFT}} = \boldsymbol{Y}_{\mathrm{ani}}^{\mathrm{ML}} + \boldsymbol{\varepsilon} = \mathbf{X}\mathbf{w} + \boldsymbol{\varepsilon}$, where $\mathbf{X}$ denotes the design matrix constructed from the pair-resolved descriptors and $\mathbf{w}$ collects all regression coefficients $w_{\mu}^{(t)}$. The residual term is then modeled as Gaussian, $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0},\, \tau^2\mathbf{I})$, which does not imply that the true residuals strictly follow a Gaussian distribution, but rather provides a working assumption conditional on the adopted linear pair-additive model, enabling a tractable quantification of uncertainty. This assumption leads to the Gaussian likelihood

$$p(\boldsymbol{Y}_{\mathrm{ani}}^{\mathrm{DFT}} \mid \mathbf{X}, \mathbf{w}) = \mathcal{N}(\mathbf{X}\mathbf{w},\, \tau^2\mathbf{I}). \tag{8}$$

A zero-mean isotropic Gaussian prior is imposed on the regression coefficients, $p(\mathbf{w}) = \mathcal{N}(\mathbf{0},\, (\tau^2/\alpha)\mathbf{I})$. Under this formulation, the posterior distribution of the coefficients is also Gaussian [63],

$$p(\mathbf{w} \mid \boldsymbol{Y}_{\mathrm{ani}}^{\mathrm{DFT}}, \mathbf{X}) = \frac{p(\boldsymbol{Y}_{\mathrm{ani}}^{\mathrm{DFT}} \mid \mathbf{X}, \mathbf{w})\, p(\mathbf{w})}{p(\boldsymbol{Y}_{\mathrm{ani}}^{\mathrm{DFT}} \mid \mathbf{X})} \tag{9}$$

$$= \mathcal{N}(\mathbf{w}_{\mathrm{MAP}},\, \boldsymbol{\Sigma}_w)\,. \tag{10}$$

with the posterior covariance given by

$$\boldsymbol{\Sigma}_w = \tau^2 \left(\mathbf{X}^\top \mathbf{X} + \alpha \mathbf{I}\right)^{-1}. \tag{11}$$

With the choice of prior $p(\mathbf{w}) = \mathcal{N}(\mathbf{0},\, (\tau^2/\alpha)\mathbf{I})$, maximizing the posterior is exactly equivalent to solving the ridge regression problem employed above. Accordingly, the posterior mean $\mathbf{w}_{\mathrm{MAP}}$ exactly coincides with the solution of standard ridge regression with regularization parameter α. In practice, the noise variance τ^2 is estimated from the residuals of the fitted model and therefore reflects intrinsic model mismatch and finite-data effects, rather than numerical noise in the underlying DFT calculations. This estimation is carried out for the ridge regression with $\alpha = 30$ used throughout this work, and we perform this Bayesian ridge inference separately for each target. Propagating the corresponding coefficient-level uncertainties [(11)] through (4) yields the shaded credible bands of $K_Y^{(t)}(d)$ shown in Figs. 4(a, c, e), corresponding to a two–standard-deviation uncertainty. Although the model does not yet perfectly reproduce the anisotropy coefficients inferred from the underlying DFT data, owing to residual uncertainty and model limitations, these credible bands indicate that the dominant trends and relative signs of the channels are captured in a statistically robust manner under the assumed linear pair-additive model, without implying a complete error estimate including systematic effects from neglected higher-order interactions.

To bridge the geometric parameters $(V, c/a)$ with the underlying atom-pair statistics, we perform the following idealized geometry analysis. In this analysis, E_{MCA}, $\mathcal{V}$, and $\mathcal{B}$ are analytically reconstructed from the learned anisotropy coefficient functions $K_Y^{(t)}(d)$ for a given $(V, c/a)$. The volume V is defined per 1×1×1 bcc unit, and (a, c) are related to $(V, c/a)$ through $a = \left(\frac{V}{c/a}\right)^{1/3}$ and $c = \left((c/a)^2 V\right)^{1/3}$.

For each orientation channel t, we evaluate the corresponding bond length $d_t(a, c)$ and the angular factor $f_t(a, c) \equiv (\sin^2\theta_x - \sin^2\theta_z)_t$, both determined purely by

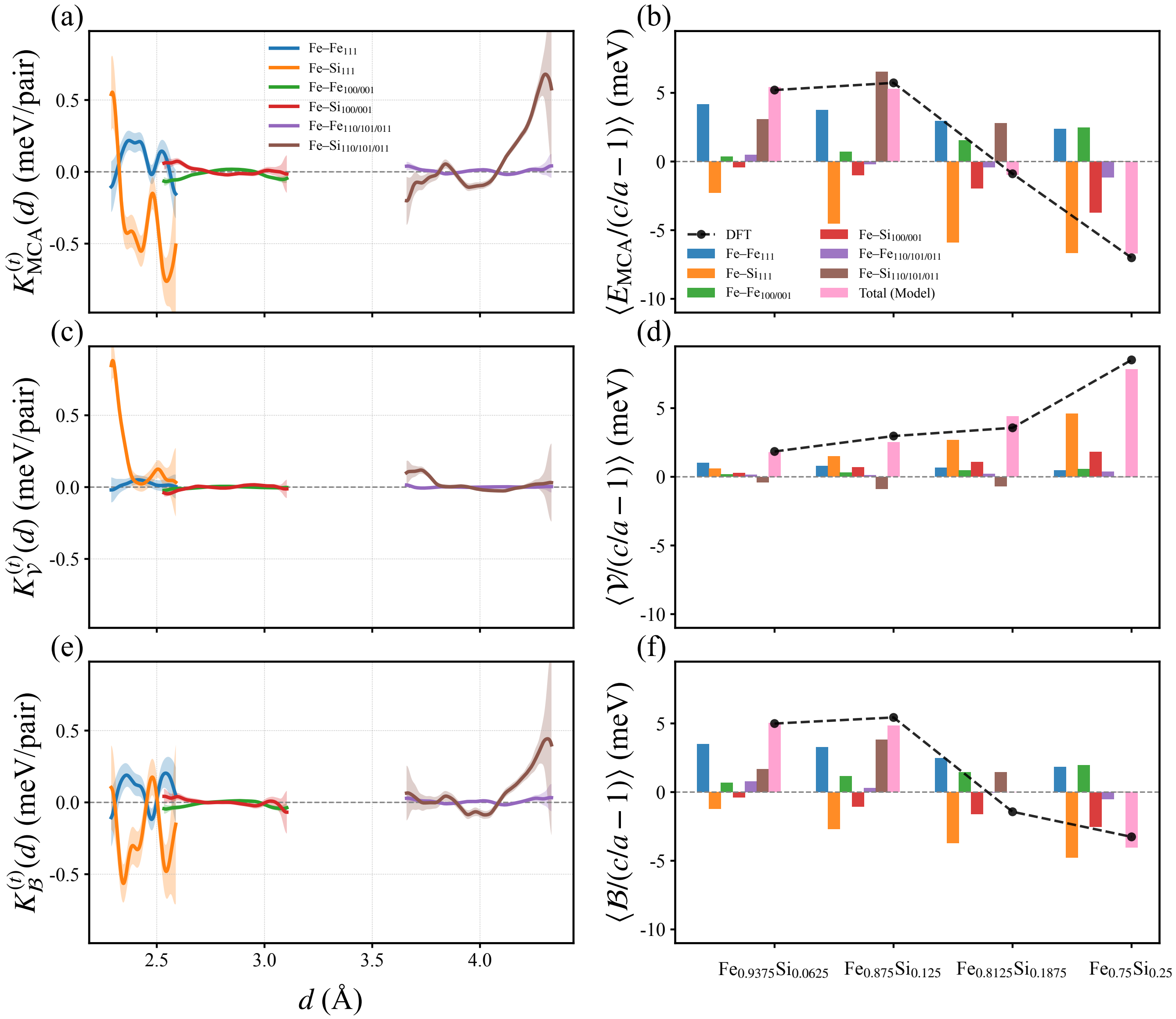


Fig. 4. Atom-pair-resolved contributions for three targets under tetragonal strain. (a, c, e) Learned anisotropy coefficient functions $K^{(t)}_{\mathrm{MCA}}(d)$, $K^{(t)}_{\mathcal{V}}(d)$, and $K^{(t)}_{\mathcal{B}}(d)$ from ridge regression for E_{MCA}, the van der Laan term $\mathcal{V}$, and the Bruno term $\mathcal{B}$, shown for Fe–Fe and Fe–Si pairs in six symmetry-aware channels (see Section 2.2). Solid curves are evaluated from Eqs. (4)–(6) using ridge regression coefficients restored to the original (unstandardized) feature scale. Shaded regions indicate posterior credible bands of $K^{(t)}_{Y}(d)$, corresponding to a two–standard-deviation uncertainty inferred from the Bayesian posterior distribution of the regression coefficients. (b, d, f) Channel-resolved average contributions to $\langle E_{\mathrm{MCA}}/(c/a-1)\rangle$, $\langle \mathcal{V}/(c/a-1)\rangle$, and $\langle \mathcal{B}/(c/a-1)\rangle$, computed by averaging over all strained states for each composition. The first six bars decompose the contributions by orientation channel, while the seventh (pink) bar shows the total model prediction (ridge fit without an intercept term). Black dots denote the corresponding DFT averages for direct comparison. Bar-plot values are reported in meV for a 2×2×2 supercell containing 16 atoms.

lattice geometry as

$$
\begin{aligned}
d_{111} &= \tfrac{1}{2}\sqrt{2a^2+c^2}, & f_{111} &= \frac{c^2-a^2}{2a^2+c^2},\\
d_{100} &= a, & f_{100} &= -1,\\
d_{001} &= c, & f_{001} &= 1,\\
d_{110} &= \sqrt{2}\,a, & f_{110} &= -\tfrac{1}{2},\\
d_{101} &= \sqrt{a^2+c^2}, & f_{101} &= \frac{c^2-a^2}{a^2+c^2},\\
d_{011} &= \sqrt{a^2+c^2}, & f_{011} &= \frac{c^2}{a^2+c^2}.
\end{aligned}
\tag{12}
$$

Each channel thus provides an analytic feature pair (d_t, f_t) that serves as the direct geometric input to the learned kernels $K^{(t)}_{Y}(d)$. For a given anisotropic target, the reconstructed response is

$$
Y_{\mathrm{ani}}(V, c/a) = \sum_t N_t\, K^{(t)}_{Y}\big[d_t(a,c)\big]\, f_t(a,c) \tag{13}
$$

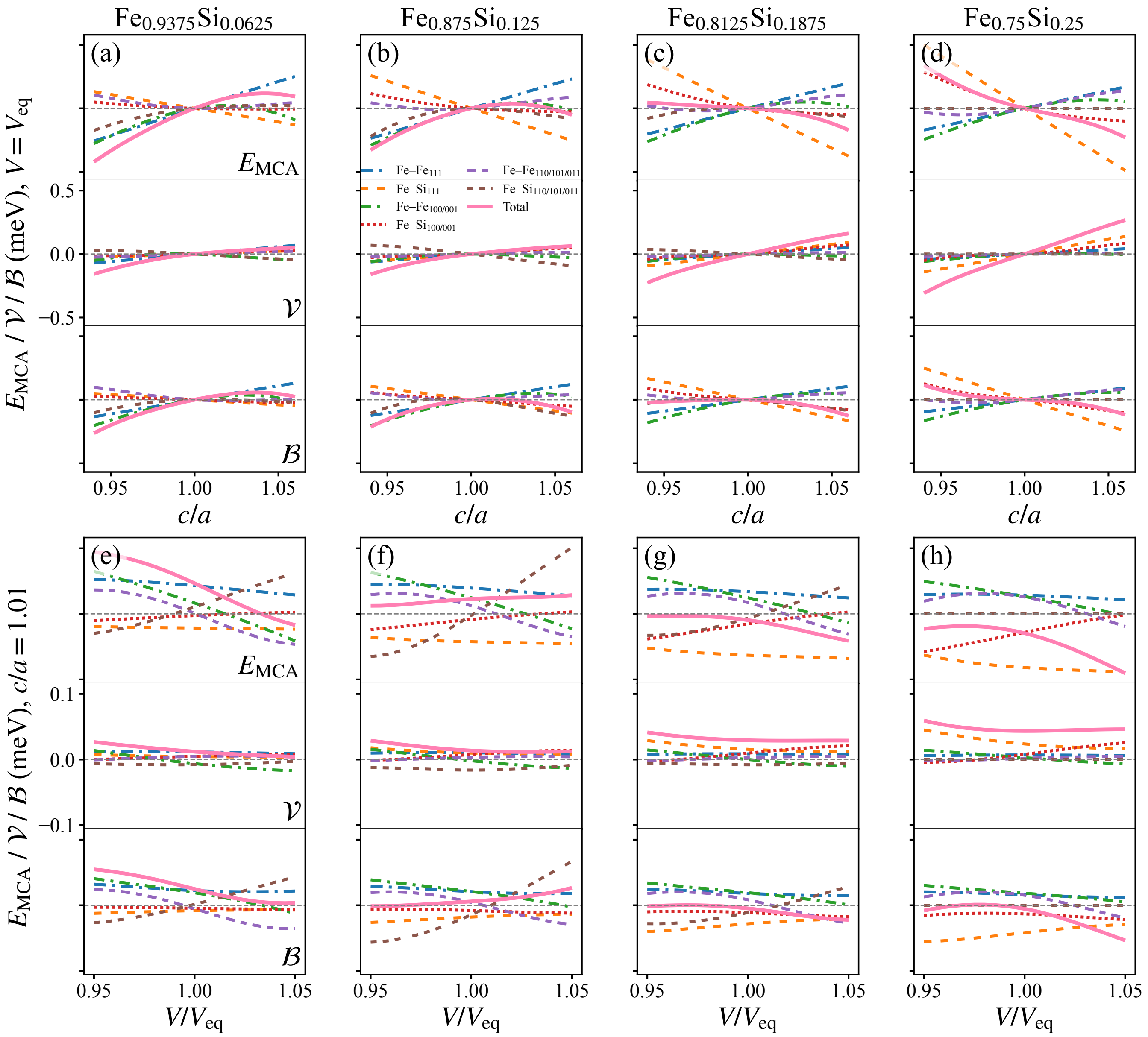


Fig. 5. Atom-pair-resolved contributions from the pair-based ridge regression model, evaluated using the idealized bcc-geometry analysis described in (12) and (13). Each panel shows three stacked bands (top to bottom) for E_{MCA}, the van der Laan term $\mathcal{V}$, and the Bruno term $\mathcal{B}$. Panels (a)–(d) scan c/a at fixed $V = V_{\mathrm{eq}}$, whereas panels (e)–(h) scan V/V_{eq} at fixed $c/a = 1.01$. For each composition and target, the curves report Fe–Fe and Fe–Si contributions in the 111 (1NN), 100/001 (2NN), and 110/101/011 (3NN) channels, together with their sum (pink); the plotted total excludes the intercept for clarity. Calculations are based on the 2×2×2 (16 atom) cells. For $\mathrm{Fe}_{1-x}\mathrm{Si}_x$ with $x = 0.0625$, 0.125, 0.1875, and 0.25, the corresponding equilibrium volumes are $V_{\mathrm{eq}} = 22.62$, 22.36, 22.23, and 21.95 Å^3 per 1×1×1 bcc unit.

where N_t are the multiplicities of each pair type and orientation channel, taken from Table 2, which reflects the actual number of equivalent pairs within a 2×2×2 supercell (16 atoms). The formulation enables a direct evaluation of orientation-resolved anisotropy responses over the broad range of tetragonal distortions (V, c/a) sampled in this work.

Fig. 5 presents the results of the pair-based ridge regeesion model with the idealized atom pair simulations described in (12) and (13). In this idealized simulation, we used a broader Gaussian width of $\sigma = 0.06$ instead of the training value $\sigma = 0.02$, in order to suppress minor numerical oscillations of $K_{\mathrm{Y}}^{(t)}$ without modifying the underlying model parameters. To connect the model analysis with physically stable states, we calculate the equilibrium volumes V_{eq} for the four compositions as those minimizing the DFT total energy at c/a=1. The values are V_{eq}=22.62, 22.36, 22.23, and 21.95 Å^3 (per 1×1×1 bcc unit) for $\mathrm{Fe}_{0.9375}\mathrm{Si}_{0.0625}$, $\mathrm{Fe}_{0.875}\mathrm{Si}_{0.125}$, $\mathrm{Fe}_{0.8125}\mathrm{Si}_{0.1875}$, and $\mathrm{Fe}_{0.75}\mathrm{Si}_{0.25}$, respectively. Figs. 5(a)–(d) show the strain response at fixed V_{eq} as c/a varies around unity, while panels (e)–(h) show the response to volume variation at fixed c/a=1.01, with three stacked bands for E_{MCA} (top), $\mathcal{V}$ (middle), and $\mathcal{B}$ (bottom). In the E_{MCA} band, it is noteworthy that when c/a is kept constant, E_{MCA} exhibits a non-negligible variation with volume. A pair-resolved analysis further shows that while the net contribution from 1NN pairs is relatively insensitive to volume, the individual 2NN and 3NN pair channels display a much stronger volume dependence and can even undergo sign reversals. In particular, for the 6.25 at.% Si composition, the sign of the total E_{MCA} can even reverse as V changes. We further note that $\mathcal{V}$ tends to increase with Si concentration, while $\mathcal{B}$ follows a trend similar to the total E_{MCA}. Although perturbative decomposition suggests $E_{\mathrm{MCA}} \approx \mathcal{B}+\mathcal{V}$, this approximation becomes progressively less accurate as Si concentration increases; for higher Si concentration, $\mathcal{B}+\mathcal{V}$ can deviate substantially from E_{MCA} and may even have the wrong sign.

In fact, the volume-dependent behavior mentioned above is also captured by the RF regression model introduced later and trained on the $(V, c/a)$ descriptors (see Fig. 6), suggesting that the pair-based ridge regression captures a shared volume-related trend rather than relying on an incidental correlation. A more detailed comparison between ridge regression and RF regression as the baseline test results is provided in Appendix F.

The anisotropy coefficients for atom pairs $K_{Y}^{(t)}(d)$ serve as physically meaningful strain-response functions whose signs and slopes directly reflect the anisotropic target responses. From a design perspective, the overall strain dependence of E_{MCA} can be tuned by considering the contributions of individual atom-pair channels: by analyzing which channels dominate at operative pair lengths and how they respond to strain, one can adjust atomic positions, substitute chemical species, or modify the equilibrium volume so as to selectively amplify or attenuate particular channel responses, thereby rebalancing the overall magnetoelastic responses.

3.2. RF regression with global $(V, c/a)$ descriptors

As a complementary reference to the pair-based ridge framework, we constructed an RF regression model based solely on the global descriptors $(V, c/a)$ to capture the overall volume and strain dependence of E_{MCA} across multiple Fe–X alloys. The RF response surface was trained to serve as an in-sample interpolant using the DFT data of $\mathrm{Fe}_{0.9375}\mathrm{Si}_{0.0625}$, $\mathrm{Fe}_{0.875}\mathrm{Si}_{0.125}$, $\mathrm{Fe}_{0.8125}\mathrm{Si}_{0.1875}$, $\mathrm{Fe}_{0.75}\mathrm{Si}_{0.25}$, $\mathrm{Fe}_{0.75}\mathrm{Ge}_{0.25}$, $\mathrm{Fe}_{0.75}\mathrm{Al}_{0.25}$ and $\mathrm{Fe}_{0.75}\mathrm{Ga}_{0.25}$, with a nominal composition index (1–7) that distinguishes these alloy types.

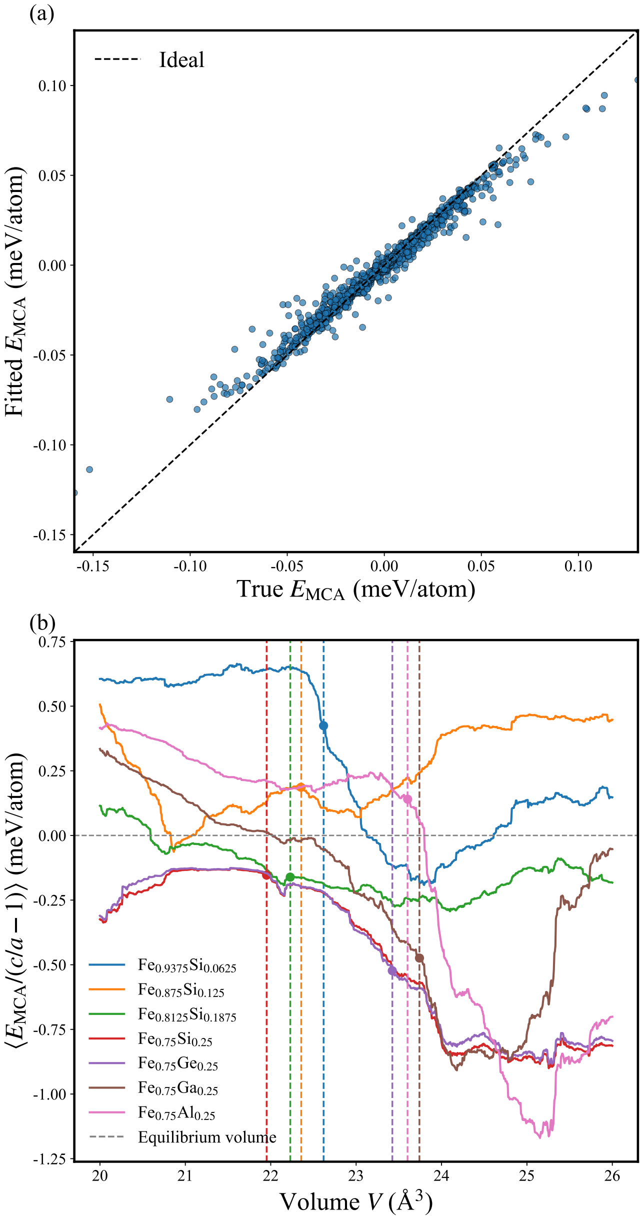


Fig. 6. Results of RF regression using the $(V, c/a)$ descriptors. (a) In-sample performance of the RF surrogate trained on (V, c/a, composition index); each point corresponds to one configuration. The dashed diagonal indicates the ideal relation. The fit yields an R^2 of 0.96. (b) Composition-resolved $E_{\mathrm{MCA}}/(c/a-1)$ versus volume. For numerical stability, only states with $|c/a-1| > 0.002$ are used. For each composition, the curve is obtained by predicting E_{MCA} on a V grid and then computing $E_{\mathrm{MCA}}/(c/a-1)$, averaged over a strain window $|c/a-1| \in [0.002, 0.03]$. Dashed vertical lines indicate each composition's equilibrium volume. Here, V is given per conventional 1×1×1 bcc unit.

Fig. 6(a) presents the fitted versus true values for all $\sim 7\times200$ configurations, yielding an R^2 score of 0.96, which indicates a high degree of predictive accuracy and a close agreement between the surrogate model and the DFT reference data. As shown in Fig. 6(b), $E_{\mathrm{MCA}}/(c/a-1)$ exhibits a pronounced dependence on V across all seven alloys, and the detailed $E_{\mathrm{MCA}}/(c/a-1)$–V relations differ markedly among some compositions. As a result, the RF regression model based solely on global geometric parameters (V, c/a) lacks generalization across different alloys. This limitation directly motivated the development of the pair-based ridge framework discussed in earlier sections. Although the RF regression lacks the ability to predict unseen compositions, it still can serve as a useful baseline to examine whether the pair-based ridge regression can reproduce the composition- and volume-dependent behavior of E_{MCA}. A quantitative comparison (baseline test) between these two approaches, namely the local pair-based ridge regression and the global RF regression, is presented in Appendix F.

We further examine the composition-dependent trends in Fig. 6(b). Within the Fe–Si series, increasing Si concentration systematically shifts the equilibrium volume V_{eq} to smaller values and simultaneously reshapes the overall $E_{\mathrm{MCA}}/(c/a-1)$–V branch, leading to a gradual evolution and even sign reversal of the magnetoelastic response. On the other hand, a useful comparison arises when considering the four neighboring elements Si, Ge, Ga, and Al, which occupy adjacent positions in the periodic table and form the Fe-based D0$_3$ compounds. D0$_3$ $\mathrm{Fe}_{0.75}\mathrm{Ge}_{0.25}$ traces a nearly identical $E_{\mathrm{MCA}}/(c/a-1)$–V curve to D0$_3$ $\mathrm{Fe}_{0.75}\mathrm{Si}_{0.25}$ but at a larger equilibrium volume. This indicates that the stronger magnetoelastic response in $\mathrm{Fe}_{0.75}\mathrm{Ge}_{0.25}$ primarily originates from the larger atomic size of Ge, which expands the equilibrium volume. D0$_3$ $\mathrm{Fe}_{0.75}\mathrm{Ga}_{0.25}$ and $\mathrm{Fe}_{0.75}\mathrm{Al}_{0.25}$ exhibit a distinctly different trend of $E_{\mathrm{MCA}}/(c/a-1)$–V, indicating that chemistry-specific effects beyond a simple volume scaling become significant in these two systems.

3.3. Exchange splitting linking volume and MCA

To connect the volume dependence to a microscopic quantity, we analyze the average exchange splitting of the Fe d states,

$$\Delta\varepsilon_{\mathrm{ex}} \equiv \langle\varepsilon_{\downarrow}\rangle - \langle\varepsilon_{\uparrow}\rangle,$$

extracted from DFT calculations for each strained structure. Fig. 7(a) shows that $\Delta\varepsilon_{\mathrm{ex}}$ increases almost linearly with V for all compositions, indicating that lattice expansion generally enhances exchange splitting by reducing interatomic hybridization of 3d states. In contrast, Fig. 7(b) reveals a more intricate, composition-dependent relation between $E_{\mathrm{MCA}}/(c/a-1)$ and $\Delta\varepsilon_{\mathrm{ex}}$, which is nonlinear and exhibits more distinct curvature among different alloys. In both panels, only configurations with $|c/a-1| > 0.002$ are included to avoid numerical instabilities near the cubic limit.

In the perturbative analysis (Appendix A), for example, the spin-flip contribution contains an explicit factor $\xi^2/\Delta\varepsilon_{\mathrm{ex}}$ (with ξ the spin–orbit strength), so changes in $\Delta\varepsilon_{\mathrm{ex}}$ directly modulate that part. In addition, shifts in $\Delta\varepsilon_{\mathrm{ex}}$ are accompanied by changes in orbital-moment anisotropy and the intra-atomic magnetic dipole, reflecting the sensitivity of spin–orbit mixing and band occupancies, which also shape the MCA. More generally, as volume changes, the accompanying variation in exchange splitting reweights the relative contributions of the spin-conserving and spin-flip channels, giving rise to a nontrivial volume dependence of the MCA.

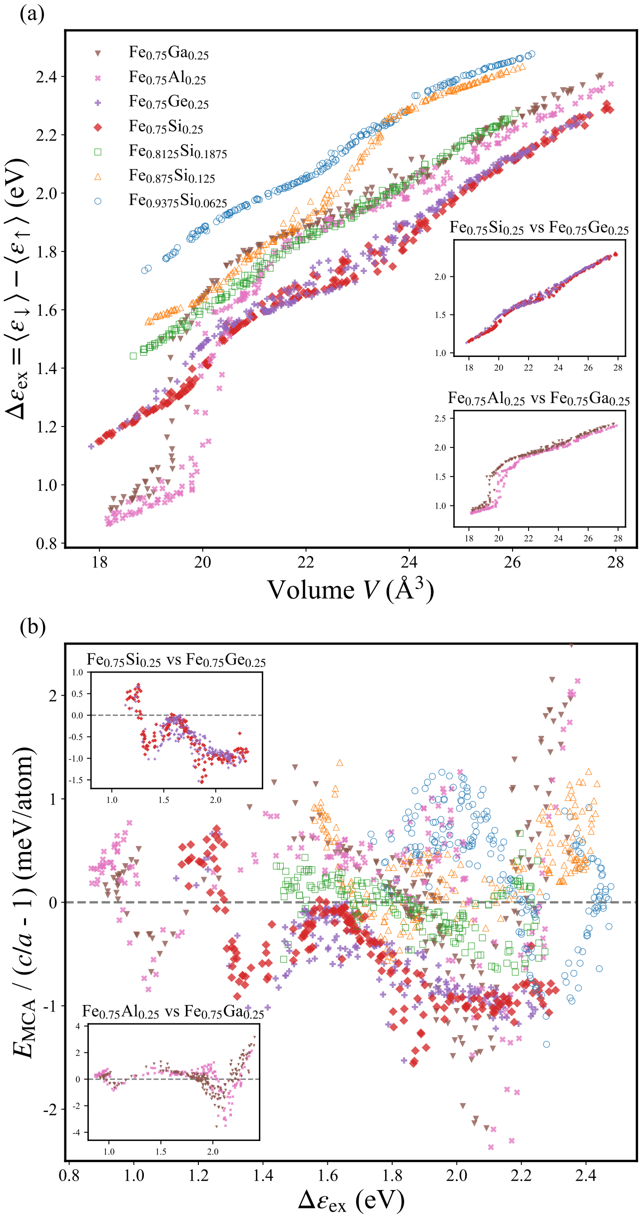


Fig. 7. Volume-dependent exchange splitting and its correlation with $E_{\mathrm{MCA}}/(c/a-1)$. (a) Average Fe 3d exchange splitting $\Delta\varepsilon_{\mathrm{ex}} = \langle\varepsilon_{\downarrow}\rangle - \langle\varepsilon_{\uparrow}\rangle$ versus volume for seven compositions. Volume V is given per conventional 1×1×1 bcc unit. (b) Correlation between $E_{\mathrm{MCA}}/(c/a-1)$ and $\Delta\varepsilon_{\mathrm{ex}}$. In both panels, only configurations with $|c/a-1| > 0.002$ are retained to avoid the numerical instability of the scaled quantity at $c/a \approx 1$.

Examining $\mathrm{Fe}_{0.75}\mathrm{Si}_{0.25}$, $\mathrm{Fe}_{0.75}\mathrm{Ge}_{0.25}$, $\mathrm{Fe}_{0.75}\mathrm{Ga}_{0.25}$, and

$Fe_{0.75}Al_{0.25}$ together allows us to investigate how closely related chemical species lead to the different magnetoelastic responses. For $Fe_{0.75}Si_{0.25}$ and $Fe_{0.75}Ge_{0.25}$, the $\Delta\varepsilon_{\mathrm{ex}}$–$V$ relations shown in Fig. 7(a) are almost indistinguishable, and the $E_{\mathrm{MCA}}/(c/a-1)$–$\Delta\varepsilon_{\mathrm{ex}}$ trends in Fig. 7(b) also coincide closely. Because both dependencies are highly similar, the resulting $E_{\mathrm{MCA}}/(c/a-1)$–V curves essentially overlap (see Fig. 6). This consistency indicates that the stronger magnetoelastic response of $Fe_{0.75}Ge_{0.25}$ compared with $Fe_{0.75}Si_{0.25}$ mainly reflects its larger equilibrium volume due to the larger atomic size of Ge, rather than any intrinsic chemical contrast with Si. For $Fe_{0.75}Al_{0.25}$ and $Fe_{0.75}Ga_{0.25}$, their $\Delta\varepsilon_{\mathrm{ex}}$–$V$ relations and $E_{\mathrm{MCA}}/(c/a-1)$–$\Delta\varepsilon_{\mathrm{ex}}$ dependencies exhibit noticeable offsets, which lead to a more pronounced separation in $E_{\mathrm{MCA}}/(c/a-1)$ across the examined volume range. This observation shows that, unlike the Si–Ge case, both volume effects and intrinsic chemical differences play important roles in shaping the magnetoelastic responses of $Fe_{0.75}Ga_{0.25}$ and $Fe_{0.75}Al_{0.25}$. Taken together with the preceding atom pair analysis, this highlights that variations in interatomic distances — and the resulting modifications of exchange splitting — play a central role in tailoring the MCA in strained Fe-based alloys.

4. Conclusion

We introduced a compact, atom-pair-centric DFT–ML framework for analyzing the strain dependence of E_{MCA} and its perturbative components $\mathcal{V}$ and $\mathcal{B}$ in Fe-rich ordered Fe–Si and related Fe–X systems. A linear pair-based ridge model trained on species- and orientation-resolved pair statistics from 2×2×2 Fe–Si cells yields smooth anisotropy coefficient functions $K_Y^{(t)}(d)$ for each target and atom-pair channel and reproduces the main DFT trends. A held-out 3×3×3 $Fe_{0.889}Si_{0.111}$ supercell is used as a direct size/composition generalization check. These coefficient functions enable distance- and orientation-resolved attribution, clarifying how composition and strain act through pair multiplicities and local coordination geometry within the adopted pair-additive surrogate. The framework suggests practical design levers such as tuning equilibrium volume through alloying and engineering local coordination to rebalance competing channels. At the same time, we emphasize that MCA originates from SOC-driven many-body electronic effects, and the present model should be viewed as an interpretable surrogate rather than a complete microscopic description.

In addition, an RF regressor trained only on global descriptors $(V, c/a)$ serves as an in-sample baseline within each composition. Its lack of generalization across different alloys highlights the value of the chemically and geometrically resolved pair-based ridge model for cross-composition modeling within the present scope. In addition to exploring various $(V, c/a)$ states for Fe–Si alloys, we further extended our analysis to several D0_3-ordered $Fe_{0.75}X_{0.25}$ systems (X = Ge, Al, Ga). Notably, the comparison between $Fe_{0.75}Si_{0.25}$ and $Fe_{0.75}Ge_{0.25}$ highlights this point: although Si and Ge are chemically similar and yield nearly overlapping $E_{\mathrm{MCA}}/(c/a-1)$–V curves, their equilibrium volumes differ. Consequently, the stronger magnetoelastic response of $Fe_{0.75}Ge_{0.25}$ can be traced primarily to its larger equilibrium volume rather than to intrinsic chemical differences. By contrast, for $Fe_{0.75}Al_{0.25}$ and $Fe_{0.75}Ga_{0.25}$, the different magnetoelastic responses arise not only from volume effects but also from intrinsic chemical differences among Al, Ga, and Si.

Overall, our results show that compact regression models based on either global macroscopic descriptors $(V, c/a)$ or local atom–pair statistics can capture strain-dependent trends in E_{MCA} for tetragonally strained bcc Fe-rich binaries within the present dataset. In particular, the pair-based ridge regression provides an interpretable decomposition into distance- and orientation-resolved pair channels, while the global-descriptor RF model offers a complementary in-sample baseline. Preliminary tests with the simple triplet descriptors did not provide a significant improvement over the present pair-based model. Future work should therefore explore more advanced higher-order descriptors and surrogate models with broader generalization. We also emphasize that interatomic-distance variations and the accompanying changes in exchange splitting are key to the design of materials with tailored MCA.

CRediT authorship contribution statement

Wen-Chuang Shang: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. Yoshihiro Gohda: Writing – review & editing, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

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

Acknowledgments

This work was partly supported by JSPS KAKENHI Grant No. JP24K01144, MEXT DXMag Grant No. JPMXP1122715503, the Spintronics Research Network of Japan (Spin-RNJ) at CSRN, the University of Osaka, and JST SPRING Grant No. JPMJSP2180. The calculations were partly carried out by using supercomputers at ISSP, the University of Tokyo, and TSUBAME4.0, Institute of Science Tokyo.

Appendix A: Second-order perturbative expression of E_{MCA}

The one-electron SOC operator is

$$\hat{H}_{\mathrm{SOC}} = \xi(\mathbf{r})\,\hat{\boldsymbol{L}}\cdot\hat{\boldsymbol{S}} = \xi(\mathbf{r})\Big(\hat{L}_z\hat{S}_z + \tfrac{1}{2}(\hat{L}_+\hat{S}_- + \hat{L}_-\hat{S}_+)\Big), \tag{A1}$$

with $\hat{L}_\pm = \hat{L}_x \pm i\hat{L}_y$ and $\hat{S}_\pm = \hat{S}_x \pm i\hat{S}_y$. Using eigenstates obtained without the SOC (unperturbed states), $\{|\phi_p^{(0)}\rangle\}$ and eigenvalues $\{\epsilon_p^{(0)}\}$, the second-order correction is given by

$$E^{(2)}_{\mathrm{SOC}} = \sum_{p\in\mathrm{occ}}\sum_{q\in\mathrm{unocc}} \frac{\left|\langle\phi_q^{(0)}|\hat{H}_{\mathrm{SOC}}|\phi_p^{(0)}\rangle\right|^2}{\epsilon_p^{(0)} - \epsilon_q^{(0)}}. \tag{A2}$$

The magnetocrystalline anisotropy is the difference

$$E^{(2)}_{\mathrm{MCA}} = E^{(2)}_{\mathrm{SOC}}([100]) - E^{(2)}_{\mathrm{SOC}}([001]). \tag{A3}$$

We write the result in an atomic orbital basis for sites I (I') with on-site SOC constants ξ_I ($\xi_{I'}$) as follows [15],

$$\begin{aligned} E^{(2)}_{\mathrm{MCA}} &= E^{\uparrow\uparrow}_{\mathrm{MCA}} + E^{\downarrow\downarrow}_{\mathrm{MCA}} + E^{\uparrow\downarrow}_{\mathrm{MCA}} + E^{\downarrow\uparrow}_{\mathrm{MCA}} &\text{(A4)}\\ &= \frac{1}{4}\sum_{I,I'}\xi_I\xi_{I'}\sum_{\ell\ell'\kappa\kappa'}\Big\{G^{\uparrow\uparrow}_{II'} + G^{\downarrow\downarrow}_{II'} - \big(G^{\uparrow\downarrow}_{II'} + G^{\downarrow\uparrow}_{II'}\big)\Big\}\\ &\quad\times\Big[\langle\ell|\hat{L}_z|\ell'\rangle\langle\kappa'|\hat{L}_z|\kappa\rangle - \langle\ell|\hat{L}_x|\ell'\rangle\langle\kappa'|\hat{L}_x|\kappa\rangle\Big], &\text{(A5)} \end{aligned}$$

where we adopt $\xi_{\mathrm{Fe}} = 0.063$ eV for Fe sites [20], ↑/↓ are used as superscripts to denote the coupling between the spin states, the first arrow denotes the spin of the occupied state and the second the spin of the unoccupied state, so that four channels appear. The band-structure kernel built from unperturbed states is

$$G^{ss'}_{II'}(\ell\ell';\kappa'\kappa) = \sum_{p\in\mathrm{occ}}\sum_{q\in\mathrm{unocc}} \frac{c^{(p)*}_{I,\ell,s}\;c^{(p)}_{I',\ell',s}\;c^{(q)*}_{I',\kappa',s'}\;c^{(q)}_{I,\kappa,s'}}{\epsilon^{(0)}_{ps} - \epsilon^{(0)}_{qs'}}. \tag{A6}$$

Here p (q) labels an occupied (unoccupied) state, ℓ (ℓ') and κ (κ') denote atomic-orbital indices, and s (s') is the spin index with $s, s' \in \{\uparrow,\downarrow\}$.

Starting from (A5), two widely used approximation forms are obtained. (i) Bruno term (denoted as $\mathcal{B}$): define the orbital-moment anisotropy on site I as

$$\Delta m^I_{\mathrm{orb}} \equiv m^I_{\mathrm{orb}}([100]) - m^I_{\mathrm{orb}}([001]) = \langle\hat{L}_x\rangle_I - \langle\hat{L}_z\rangle_I. \tag{A7}$$

One obtains Bruno's relation [14, 15, 20]:

$$E^{\downarrow\downarrow}_{\mathrm{MCA}} \;\approx\; \mathcal{B} = -\frac{1}{4\mu_B}\sum_I \xi_I\,\Delta m^I_{\mathrm{orb}}, \tag{A8}$$

which could become dominant in some typical ferromagnets, allowing the contributions of $\uparrow\uparrow$, $\uparrow\downarrow$, and $\downarrow\uparrow$ channels to be slight. (ii) van der Laan term (denoted as $\mathcal{V}$): considering the spin-flip term $\uparrow\downarrow$ of (A5), one assumes a constant average exchange splitting $\epsilon^{(0)}_{p\downarrow} - \epsilon^{(0)}_{q\uparrow} \approx \Delta\varepsilon_{\mathrm{ex}}$ and further approximates the occupied majority-spin subspace as complete, $\sum_p^{\mathrm{occ}} |p\uparrow\rangle\langle p\uparrow| \approx 1$, then one finds the relation between $E^{\uparrow\downarrow}_{\mathrm{MCA}}$ and the intra-atomic magnetic dipole (related to spin-density quadrupole) moment [15, 16, 20]:

$$E^{\uparrow\downarrow}_{\mathrm{MCA}} \;\approx\; \mathcal{V} = \frac{21}{8\mu_B}\sum_I \frac{\xi_I^2}{\Delta\varepsilon_{\mathrm{ex}}}\, m^I_{\mathrm{T}}, \tag{A9}$$

here

$$m^I_{\mathrm{T}} \equiv -\Big\langle \hat{Q}_{zz}\,\hat{S}_z\Big\rangle_I \frac{\mu_B}{\hbar}, \tag{A10a}$$

$$\hat{Q}_{zz} \equiv \frac{2}{21}\big(3\hat{L}_z^{\,2} - \hat{L}^{\,2}\big). \tag{A10b}$$

Under the validity of the above approximations involved, ignoring $E^{\uparrow\uparrow}_{\mathrm{MCA}}$ and $E^{\downarrow\uparrow}_{\mathrm{MCA}}$ yields

$$E_{\mathrm{MCA}} \;\approx\; \mathcal{B} + \mathcal{V}. \tag{A11}$$

Nevertheless, these approximations are not universally reliable and may break down in certain materials.

Appendix B: Symmetry-consistency checks

We verified the two symmetry constraints built into the model. For each structure, we apply the symmetry operations to the bond-level angular information, rebuild the transformed descriptor vectors $\mathbf{X}_{\mathrm{swap}}$ and $\mathbf{X}_{\mathrm{mirror}}$, and then re-evaluate them with the same trained model to obtain the corresponding predictions $\hat{\mathbf{Y}}_{\mathrm{ani,swap}}$ and $\hat{\mathbf{Y}}_{\mathrm{ani,mirror}}$ from the original prediction vector $\hat{\mathbf{Y}}_{\mathrm{ani}}$. In practice, this means that the pair distances and pair-channel assignments are kept fixed, while the angular factor $\sin^2\theta_x - \sin^2\theta_z$ is rebuilt after the corresponding coordinate transformation. No coefficients are refitted in this test: the transformed descriptors are passed through the deployed model exactly as in the original evaluation. The check therefore probes whether the descriptor definition itself and the trained linear mapping are jointly consistent with the required symmetry operations, rather than whether a separately retrained model could recover the same behavior. Both the training structures and the held-out 3×3×3 test structures are included in this verification.

For [100] ↔ [001] exchange, the anisotropy target is odd, so the expected relation is $\hat{\mathbf{Y}}_{\mathrm{ani,swap}} = -\hat{\mathbf{Y}}_{\mathrm{ani}}$. At descriptor level this gives

$$\Delta\mathbf{X}_{\mathrm{swap}} = \mathbf{X}_{\mathrm{swap}} + \mathbf{X} = \mathbf{0}, \tag{B1}$$

$$\Delta\hat{\mathbf{Y}}_{\mathrm{ani,swap}} = \hat{\mathbf{Y}}_{\mathrm{ani,swap}} + \hat{\mathbf{Y}}_{\mathrm{ani}} = \mathbf{0}. \tag{B2}$$

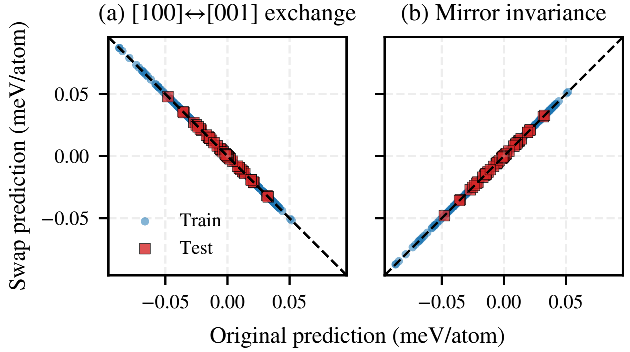


Fig. 8. Symmetry-consistency checks for the deployed pair-based ridge model. (a) [100] ↔ [001] exchange, with expected $\hat{\mathbf{Y}}_{\mathrm{ani,swap}} = -\hat{\mathbf{Y}}_{\mathrm{ani}}$. (b) mirror-$x$, with expected $\hat{\mathbf{Y}}_{\mathrm{ani,mirror}} = \hat{\mathbf{Y}}_{\mathrm{ani}}$. The strict criteria are $\Delta\mathbf{X}_{\mathrm{swap}} = \Delta\mathbf{X}_{\mathrm{mirror}} = \Delta\hat{\mathbf{Y}}_{\mathrm{ani,swap}} = \Delta\hat{\mathbf{Y}}_{\mathrm{ani,mirror}} = \mathbf{0}$, which are satisfied within numerical precision.

For mirror-x, the anisotropy target is even, so the expected relation is $\hat{\mathbf{Y}}_{\mathrm{ani,mirror}} = \hat{\mathbf{Y}}_{\mathrm{ani}}$, giving

$$\Delta\mathbf{X}_{\mathrm{mirror}} = \mathbf{X}_{\mathrm{mirror}} - \mathbf{X} = \mathbf{0}, \tag{B3}$$

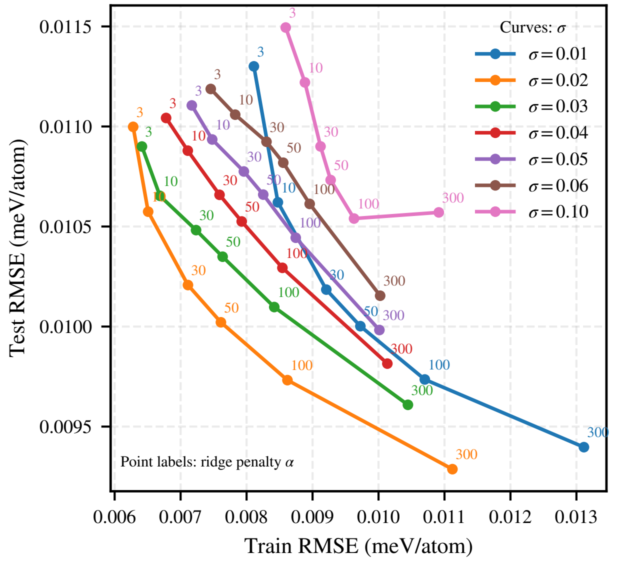
$$\Delta\hat{\mathbf{Y}}_{\mathrm{ani,mirror}} = \hat{\mathbf{Y}}_{\mathrm{ani,mirror}} - \hat{\mathbf{Y}}_{\mathrm{ani}} = \mathbf{0}. \tag{B4}$$

Numerically, both operations satisfy the strict criteria above for all training and held-out structures within reported precision. As shown in Fig. 8, the transformed predictions fall exactly on the expected odd and even symmetry lines.

Appendix C: Hyperparameter robustness

To visualize sensitivity to descriptor resolution and regularization, we scanned the RBF width σ, ridge penalty α, and number of centers N_μ around the baseline setting. Fig. 9 summarizes the resulting test RMSE. Panel (a) varies σ and α at fixed $N_\mu = 20$, while panel (b) varies N_μ at $\sigma = 0.02$ and 0.05 over the same range of α. The baseline setting $(N_\mu, \sigma, \alpha) = (20, 0.02\ \text{Å}, 30)$ lies in a stable low-error region rather than at an isolated sharp optimum. In particular, smaller σ values can sharpen the apparent in-sample fit, but they also make the result more sensitive to regularization and to noise-like fluctuations. These trends support the use of the baseline hyperparameters as a balanced working point for the present dataset. The scan also clarifies the distinct roles of the three hyperparameters. The width σ controls how locally the descriptor resolves strain-induced bond-length shifts, the ridge penalty α suppresses unstable coefficient variations, and N_μ sets the radial resolution of the coefficient curves. Panel (a) shows that the smallest σ values are not uniformly preferable: they can reduce the train error, but only at the cost of stronger sensitivity to α and a narrower stable region. Panel (b) indicates that the gain of increasing N_μ beyond the baseline is modest.

(a) Sigma scan at $N_\mu = 20$



(b) Center-count scan at $\sigma = 0.02$ and 0.05

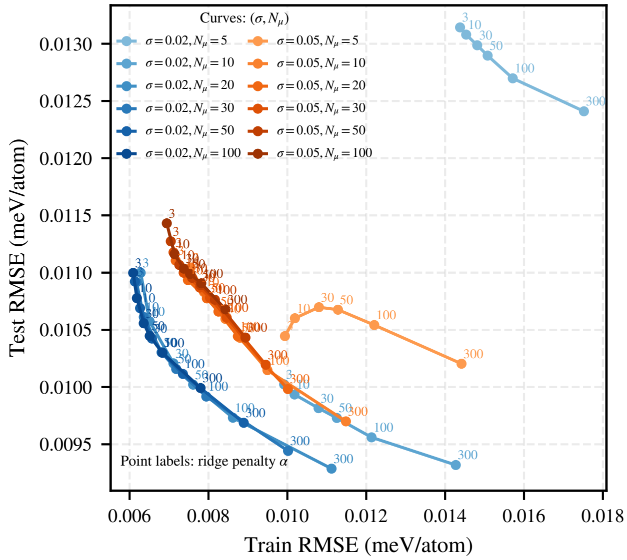


Fig. 9. Hyperparameter robustness as a trade-off map for the pair-based ridge model. (a) Kernel-width scan over (σ, α) at fixed $N_\mu = 20$. (b) Center-count scan over N_μ at $\sigma = 0.02$ and 0.05. Point labels indicate the ridge regularization strength. The baseline setting $(N_\mu, \sigma, \alpha) = (20, 0.02\ \text{Å}, 30)$ lies in a stable low-error region rather than at an isolated sharp optimum.

Appendix D: Neighbor-shell ablation

To test the descriptor truncation level, we compared three pair sets: 1NN only, 1NN+2NN, and 1NN+2NN+3NN. For each truncation level, the model

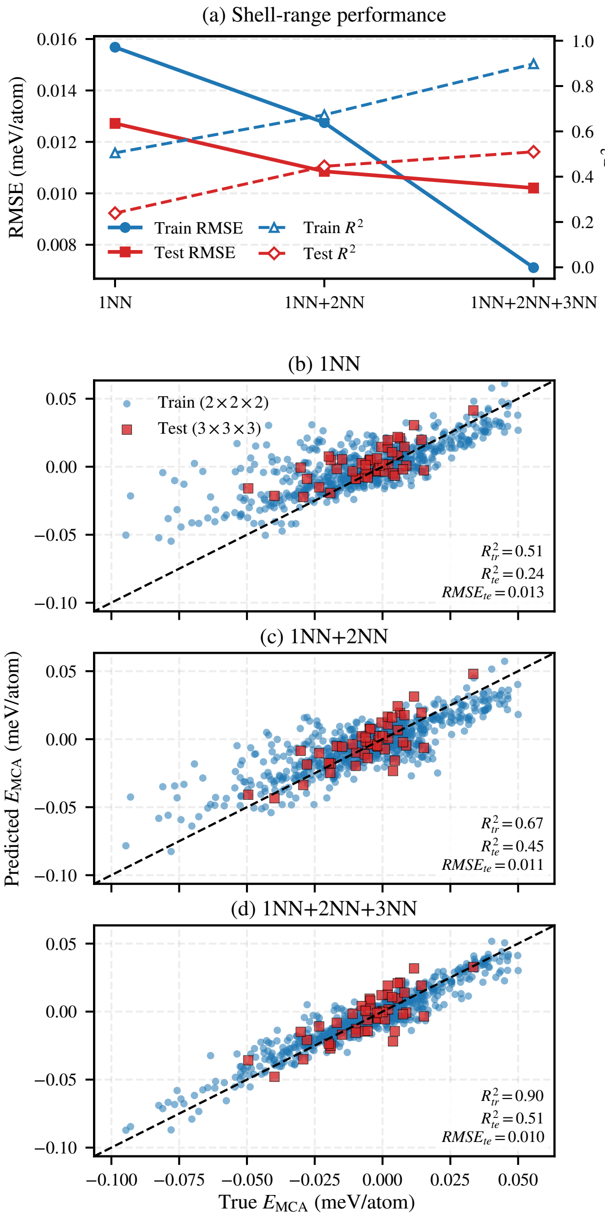


Fig. 10. Neighbor-shell ablation for the pair-based ridge model. (a) Train/test RMSE and R^2 versus shell range. (b)–(d) Predicted-versus-true results for 1NN, 1NN+2NN, and 1NN+2NN+3NN.

was trained on the same 2×2×2 compositions and evaluated on the same held-out 3×3×3 $\mathrm{Fe}_{0.889}\mathrm{Si}_{0.111}$ set, using identical ridge regularization. The descriptor dimensionality increases from 40 to 80 to 120 features.

The trend in Fig. 10 is systematic. From 1NN to 1NN+2NN+3NN, the train RMSE decreases from 0.016 to 0.013 to 0.0071 meV/atom, the test RMSE decreases from 0.013 to 0.011 to 0.010 meV/atom, and the test R^2 improves from 0.24 to 0.45 to 0.51. Within the present Fe-rich ordered Fe–Si dataset, including pairs up to 3NN therefore yields a clear gain and nearly saturates the present model class.

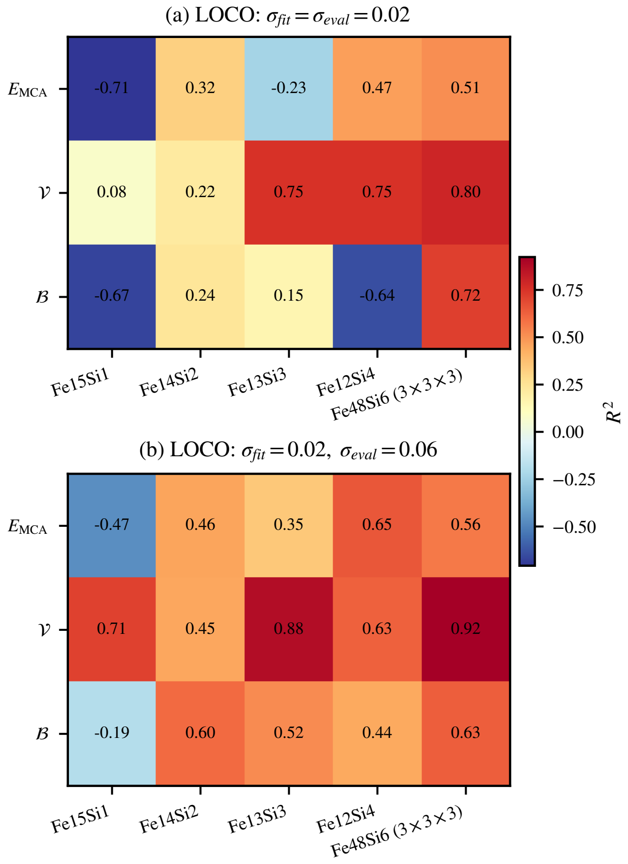


Fig. 11. LOCO validation, where each composition is excluded in turn from training and used as the test composition. Heat maps show the test R^2 values for E_{MCA}, $\mathcal{V}$, and $\mathcal{B}$. (a) The evaluation with $\sigma_{\mathrm{fit}} = \sigma_{\mathrm{eval}} = 0.02$. (b) Cross-$\sigma$ evaluation with $\sigma_{\mathrm{fit}} = 0.02$ and a broader evaluation kernel, $\sigma_{\mathrm{eval}} = 0.06$. Columns correspond to the held-out compositions, and rows correspond to the target quantities. The cell colors and overlaid numbers indicate the corresponding test R^2 values.

Appendix E: Composition-wise Leave-One-Composition-Out validation

To probe composition-wise generalization beyond the single held-out 3×3×3 test cell, we performed LOCO validation across the four 2×2×2 Fe–Si compositions and also report the external-cell evaluation. Fig. 11 summarizes the target-wise test R^2 values for E_{MCA}, $\mathcal{V}$, and $\mathcal{B}$. Panel (a) uses the standard setting $\sigma_{\mathrm{fit}} = \sigma_{\mathrm{eval}} = 0.02$, whereas panel (b) uses a cross-σ evaluation with $\sigma_{\mathrm{fit}} = 0.02$ and $\sigma_{\mathrm{eval}} = 0.06$. The cross-$\sigma$ protocol improves or stabilizes the held-out behavior for most folds and targets, consistent with the interpretation that a slightly broader evaluation kernel suppresses noise-sensitive variations while retaining the fitted pair-level trends.

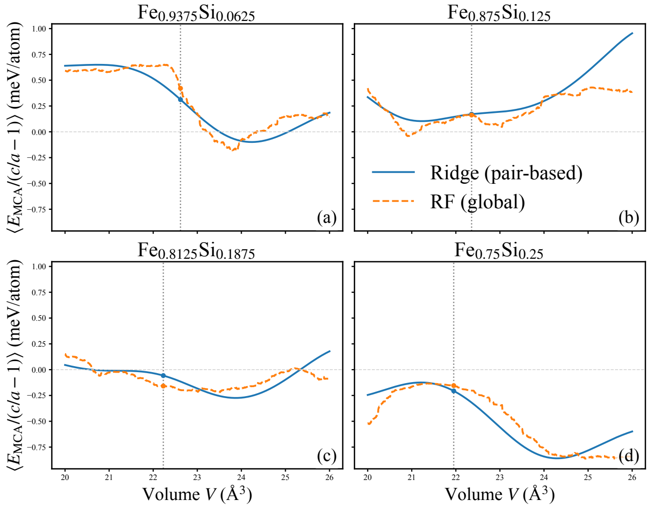


Fig. 12. Comparison between the pair-based ridge model (solid) and RF model (dashed) for four Fe–Si compositions. Vertical dotted lines mark the equilibrium volumes V_{eq}.

Appendix F: Comparison between ridge and RF regression models

We further compared the pair-based ridge model with an RF model trained on the same Fe–Si dataset. The ridge model encodes the V and c/a dependence through the idealized bcc geometry of (12) and (13) together with the pair summation of (2) and (3), whereas the RF model uses V and c/a directly as global inputs. Both models were evaluated on a uniform $(V, c/a)$ grid, and $\langle E_{\mathrm{MCA}}/(c/a-1)\rangle$ was averaged over $|c/a-1| \in [0.002, 0.03]$. Fig. 12 compares the resulting curves for four Fe–Si compositions. The ridge model reproduces the overall magnitude, sign evolution, and composition dependence captured by the RF model, while deviations emerge far from equilibrium volumes, where the additivity approximation and the coarse orientation-channel treatment become less accurate. This comparison supports the physical relevance of the pair-based description while clarifying its present limitations.

Data availability

The data and code used in this study can be found at the ISSP repository via following link. `https://isspns-gitlab.issp.u-tokyo.ac.jp/shangwenchuang/pair-based-ridge-regression-model`

---

[1] J. Stöhr, J. Magn. Magn. Mater. **200**, 470 (1999).
[2] R. Wu and A. J. Freeman, J. Magn. Magn. Mater. **200**, 498 (1999).
[3] R. Grössinger, R. S. Turtelli, and N. Mehmood, IOP Conf. Ser.: Mater. Sci. Eng. **60**, 012002 (2014).
[4] C. Mudivarthi, M. Laver, J. Cullen, A. B. Flatau, and M. Wuttig, J. Appl. Phys. **107**, 09A957 (2010).
[5] H. Wang, Y. N. Zhang, R. Q. Wu, L. Z. Sun, D. S. Xu, and Z. D. Zhang, Sci. Rep. **3**, 3521 (2013).
[6] S. Fujii, T. Usami, Y. Shiratsuchi, A. M. Kerrigan, A. M. Yatmeidhy, S. Yamada, T. Kanashima, R. Nakatani, V. K. Lazarov, T. Oguchi, *et al.*, NPG Asia Mater. **14**, 43 (2022).
[7] J. Atulasimha and A. B. Flatau, Smart Mater. Struct. **20**, 043001 (2011).
[8] D. Odkhuu, W. S. Yun, and S. C. Hong, J. Appl. Phys. **111**, 063911 (2012).
[9] W. J. Carr and R. Smoluchowski, Phys. Rev. **83**, 1236 (1951).
[10] Q. Xing, D. Wu, and T. A. Lograsso, J. Appl. Phys. **107**, 09A911 (2010).
[11] E. Adams, J. Appl. Phys. **33**, 1214 (1962).
[12] E. Callen and H. B. Callen, Phys. Rev. **139**, A455 (1965).
[13] Y. Zhang and R. Wu, IEEE Trans. Magn. **47**, 4044 (2011).
[14] P. Bruno, Phys. Rev. B **39**, 865 (1989).
[15] Y. Miura and J. Okabayashi, J. Phys.: Condens. Matter **34**, 473001 (2022).
[16] G. van der Laan, J. Phys.: Condens. Matter **10**, 3239 (1998).
[17] J. R. Cullen, A. E. Clark, J. B. Restorff, and K. B. Hathaway, J. Magn. Magn. Mater. **226**, 948 (2000).
[18] J. X. Cao, Y. N. Zhang, W. J. Ouyang, and R. Q. Wu, Phys. Rev. B **80**, 104414 (2009).
[19] L. Zheng, C.-B. Jiang, J.-X. Shang, and H.-B. Xu, Chin. Phys. B **18**, 1647 (2009).
[20] A. M. Yatmeidhy and Y. Gohda, Appl. Phys. Express **16**, 053001 (2023).
[21] A. M. Yatmeidhy and Y. Gohda, Sci. Technol. Adv. Mater. **25**, 2391268 (2024).
[22] Y. Hamazaki and Y. Gohda, J. Appl. Phys. **126**, 233902 (2019).
[23] R. Costa-Amaral and Y. Gohda, Phys. Rev. Applied **15**, 064014 (2021).
[24] Z. Torbatian, T. Ozaki, S. Tsuneyuki, and Y. Gohda, Appl. Phys. Lett. **104**, 242403 (2014).
[25] S. Nakamura and Y. Gohda, Phys. Rev. B **96**, 245416 (2017).
[26] J. Behler and M. Parrinello, Phys. Rev. Lett. **98**, 146401 (2007).
[27] M. Zhang, K. Hibi, and J. Inoue, Phys. Rev. B **110**, 054110 (2024).
[28] K. T. Schütt, H. E. Sauceda, P.-J. Kindermans, A. Tkatchenko, and K.-R. Müller, J. Chem. Phys. **148**, 241722 (2018).
[29] T. Xie and J. C. Grossman, Phys. Rev. Lett. **120**, 145301 (2018).
[30] K. Choudhary and B. DeCost, npj Comput. Mater. **7**, 185 (2021).
[31] T. Liao, W. Xia, M. Sakurai, R. Wang, C. Zhang, H. Sun, K.-M. Ho, C.-Z. Wang, and J. R. Chelikowsky, Phys. Rev. Mater. **6**, 024402 (2022).
[32] J. Navrátil, R. Topolnicki, M. Otyepka, P. Błoński, *et al.*, npj Comput. Mater. **11**, 138 (2025).
[33] P. Singh, Adv. Funct. Mater. , e25433 (2025).
[34] S. Cui and I.-H. Jung, CALPHAD **56**, 108 (2017).
[35] M. Bernhard, D. Kavic, P. Presoly, T.-G. Wi, W.-B. Park, R. Rössler, A. Jungreithmeier, S. Ilie, C. Bernhard, and Y.-B. Kang, Metall. Mater. Trans. B **56**, 2249 (2025).
[36] R. A. Khalaniya and A. V. Shevelkov, J. Solid State Chem. **270**, 118 (2019).
[37] H. Okamoto, J. Phase Equilib. Diffus. **29**, 292 (2008).
[38] T. Zienert and O. Fabrichnaya, J. Alloys Compd. **743**, 795 (2018).
[39] A. T. Phan, M.-K. Paek, and Y.-B. Kang, Acta Mater. **79**, 1 (2014).
[40] A. K. Mohamed, V. V. Palacheva, V. V. Cheverikin, E. N. Zanaeva, W.-C. Cheng, V. Kulitckii, S. Divinski, G. Wilde, and I. S. Golovin, J. Alloys Compd. **846**, 156486 (2020).
[41] K. Han, M. Saito, J. Xia, I. Ohnuma, and R. Kainuma, J. Alloys Compd. **919**, 165810 (2022).
[42] I. Ohnuma, S. Abe, S. Shimenouchi, T. Omori, R. Kainuma, and K. Ishida, ISIJ Int. **52**, 540 (2012).
[43] K. Ohno, R. Kuwahara, R. Sahara, T. N. Pham, S. Bhattacharyya, Y. Kawazoe, and K. Fujisaki, ISIJ Int. **63**, 553 (2023).
[44] P. Villars, L. D. Calvert, and W. B. Pearson, *Pearson's Handbook of Crystallographic Data for Intermetallic Phases* (ASM International, Materials Park, OH, 1985).
[45] S. C. Hong, W. S. Yun, and R. Wu, Phys. Rev. B **79**, 054419 (2009).
[46] G. Kresse and J. Furthmüller, Comput. Mater. Sci. **6**, 15 (1996).
[47] G. Kresse and D. Joubert, Phys. Rev. B **59**, 1758 (1999).
[48] P. E. Blöchl, Phys. Rev. B **50**, 17953 (1994).
[49] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[50] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[51] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[52] H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**, 5188 (1976).
[53] G. H. O. Daalderop, P. J. Kelly, and M. F. H. Schuurmans, Phys. Rev. B **41**, 11919 (1990).
[54] X. D. Wang, R. Q. Wu, D. S. Wang, and A. J. Freeman, Phys. Rev. B **54**, 61 (1996).
[55] S. V. Halilov, A. Y. Perlov, P. M. Oppeneer, A. N. Yaresko, and V. N. Antonov, Phys. Rev. B **57**, 9557 (1998).
[56] A. E. Hoerl and R. W. Kennard, Technometrics **12**, 55 (1970).
[57] L. Breiman, Mach. Learn. **45**, 5 (2001).
[58] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, J. Mach. Learn. Res. **12**, 2825 (2011).
[59] T. Burkert, L. Nordström, O. Eriksson, and O. Heinonen, Phys. Rev. Lett. **93**, 027203 (2004).

[60] K. Barmak, J. Kim, L. H. Lewis, K. R. Coffey, M. F. Toney, A. J. Kellock, and J.-U. Thiele, J. Appl. Phys. **98**, 033904 (2005).

[61] Y. Kota and A. Sakuma, J. Phys. Soc. Jpn. **83**, 034715 (2014).

[62] A. Lisfi, C. M. Williams, L. T. Nguyen, J. C. Lodder, A. Coleman, H. Corcoran, A. Johnson, P. Chang, A. Kumar, and W. Morgan, Phys. Rev. B **76**, 054405 (2007).

[63] C. M. Bishop, *Pattern Recognition and Machine Learning* (Springer, New York, NY, 2006).