Universal chemical formula dependence of *ab initio* low-energy effective Hamiltonian in single-layer carrier-doped cuprate superconductors: Study using a hierarchical dependence extraction algorithm

Jean-Baptiste Morée [1] and Ryotaro Arita [1,2]

[1] *RIKEN Center for Emergent Matter Science, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan*

[2] *Research Center for Advanced Science and Technology, University of Tokyo, Komaba, Meguro-ku, Tokyo 153-8904, Japan*

(Received 11 March 2024; revised 3 June 2024; accepted 5 June 2024; published 2 July 2024)

We explore the possibility to control the superconducting transition temperature at optimal hole doping $T_c^{\rm opt}$ in cuprates by tuning the chemical formula (CF). $T_c^{\rm opt}$ can be theoretically predicted from the parameters of the *ab initio* low-energy effective Hamiltonian with one antibonding (AB) Cu3$d_{x^2-y^2}$/O2p_σ orbital per Cu atom in the CuO_2 plane, notably the nearest-neighbor hopping amplitude $|t_1|$ and the ratio $u = U/|t_1|$, where U is the onsite effective Coulomb repulsion. However, the CF dependence of $|t_1|$ and u is a highly nontrivial question. In this paper, we propose the universal dependence of $|t_1|$ and u on the CF and structural features in hole doped cuprates with a single CuO_2 layer sandwiched between block layers. To do so, we perform extensive *ab initio* calculations of $|t_1|$ and u and analyze the results by employing a machine-learning method called hierarchical dependence extraction (HDE). The main results are (a) $|t_1|$ has a main-order dependence on the radii $R_{\rm X}$ and $R_{\rm A}$ of the apical anion X and cation A in the block layer. ($|t_1|$ increases when $R_{\rm X}$ or $R_{\rm A}$ decreases.) (b) u has a main-order dependence on the ionic charge $Z_{\rm X}$ of X and the hole doping δ of the AB orbital. (u decreases when $|Z_{\rm X}|$ increases or δ increases.) We elucidate and discuss the microscopic mechanism of items (a) and (b). We demonstrate the predictive power of the HDE by showing the consistency between items (a) and (b) and results from previous works. The present results provide a basis for optimizing superconducting properties in cuprates and possibly akin materials. Also, the HDE method offers a general platform to identify dependencies between physical quantities.

DOI: 10.1103/PhysRevB.110.014502

I. INTRODUCTION

One of the grand challenges in condensed-matter physics is the design of superconducting materials with high transition temperature T_c. The diverse distribution of $T_c^{\rm opt}$ (the experimental T_c at optimal hole doping) in carrier-doped superconducting cuprates provides useful insights into such design. In carrier doped superconducting cuprates, we have $T_c^{\rm opt} \simeq$ 10–138 K at ambient pressure [1–13] and up to $T_c^{\rm opt} \simeq$ 166 K under pressure in $HgBa_2Ca_2Cu_3O_8$ (Hg1223) [2,3]. This diverse distribution is already present in single-layer cuprates, in which $T_c^{\rm opt} \simeq$ 10–94 K at ambient pressure and up to $T_c^{\rm opt} \simeq$ 110 K in $HgBa_2CuO_4$ under pressure [2]. Thus, single-layer carrier doped cuprates are a platform of choice to investigate the microscopic mechanism and origin of the materials dependence of $T_c^{\rm opt}$.

The diverse distribution of $T_c^{\rm opt}$ can be described by the materials dependent AB Hamiltonian parameters, especially $|t_1|$ and $u = U/|t_1|$ [14]. Indeed, the scaling

$$T_c^{\rm opt} \simeq 0.16|t_1|F_{\rm SC} \tag{1}$$

was proposed [14], in which the dimensionless superconducting order parameter $F_{\rm SC}$ mainly depends on u (see Appendix A). The u dependence of $F_{\rm SC}$ is summarized below. $F_{\rm SC}$ is zero for $u \lesssim 6.5$ and increases sharply with increasing $u \simeq$ 6.5–8.0 (weak-coupling regime), reaching a maximum at $u_{\rm opt} \simeq$ 8.0–8.5 (optimal regime); then, $F_{\rm SC}$ decreases with increasing $u \gtrsim 9.0$ (strong-coupling regime). For a given material, $|t_1|$ and u can be calculated by using the multiscale *ab initio* scheme for correlated electrons (MACE) [15–18], which allowed to establish Eq. (1).

Thus, a key point for materials design of higher-$T_c^{\rm opt}$ cuprates is to elucidate the universal chemical formula (CF) dependence of $|t_1|$ and u. In previous works on cuprates such as $HgBa_2CuO_4$, $Bi_2Sr_2CuO_6$ (Bi2201), $Bi_2Sr_2CaCu_2O_8$ (Bi2212), and $CaCuO_2$ [19] as well as Hg1223 [20], the nontrivial dependence of $|t_1|$ and u on the interatomic distances and the CF has been partly clarified. However, the more general CF dependence of $|t_1|$ and u is required to obtain a thorough understanding of the CF dependence of $T_c^{\rm opt}$.

The goal of this paper is twofold. First, we propose a machine-learning procedure that is tailor-made to extract the nonlinear dependencies of a given quantity y on other quantities x_i from the main-order to the higher-order. This procedure is denoted as hierarchical dependence extraction (HDE). Second, we propose the universal CF dependence of $|t_1|$ and u in single-layer cuprates, by performing explicit *ab initio* calculations of the AB Hamiltonian for a training set that is representative of single-layer cuprates (including copper oxides, oxychlorides, and oxyfluorides), and applying the HDE to analyze the results and construct expressions of $|t_1|$ and u. We generalize the existing MACE procedure to

obtain the crystal parameters as a function of the chemical variables (the radii and charges of the cations and anions in the block layer), and *in fine* the AB Hamiltonian as a function of the chemical variables. The combination of the generalized MACE (gMACE) procedure with the analysis of the results by the HDE is denoted as gMACE + HDE. We demonstrate the predictive power of the HDE by showing the consistency between the universal CF dependencies of $|t_1|$ and u obtained by employing the gMACE + HDE and previous results on $HgBa_2CuO_4$, Bi2201, Bi2212, $CaCuO_2$, and Hg1223 [19,20].

This paper is organized as follows: Section II gives an overview of the main-order dependence (MOD) of $|t_1|$ and u on the chemical variables. Section III describes the HDE and the gMACE methodologies employed in this paper, and how the HDE is applied to analyze the results of the gMACE calculation in the gMACE + HDE. Section IV details the results on the MOD of $|t_1|$ and u on the chemical variables and proposes the microscopic mechanism underlying to this MOD. Section V discusses the results from the perspective of Eq. (1) and proposes guidelines to optimize the value of $T_c^{\rm opt}$ in future design of single-layer cuprates for which the gMACE calculation is performed. Section VI is the conclusion. Appendix A reminds the u dependence of $F_{\rm SC}$ from Ref. [14] and justifies the focus on $|t_1|$ and u to predict the materials dependence of $T_c^{\rm opt}$. Appendix B details the choice of the training set of single-layer cuprates. Appendix C gives details on the HDE. Appendix D gives details on the gMACE and the values of the intermediary quantities obtained in the gMACE calculation. Appendix E gives the analysis of the competition between variables in the MOD. Appendix F details the robustness of the physical dependencies and mathematical expressions presented in Sec. II when the number of compounds in the training set is reduced. Appendix G gives details on the hole doping dependence of the screening for each compound in the training set, and possible implications on superconducting properties. Appendix H gives complements on the density of states near the Fermi level in hole-doped oxychlorides.

II. OVERVIEW: MAIN-ORDER DEPENDENCE OF AB HAMILTONIAN PARAMETERS ON CHEMICAL FORMULA

Here, we give an overview of the CF dependence of AB Hamiltonian parameters that is obtained by applying our theoretical scheme. (Details on the results are given later in Sec. IV, and prescriptions to optimize $T_c^{\rm opt}$ based on these results are proposed in Sec. V.) The MODs of $|t_1|$ and u are summarized below in (I) and (II) and illustrated in Fig. 1.

(I) $|t_1|$ mainly depends on the crystal ionic radii $R_{\rm X}$ and $R_{\rm A}$ of the apical anion X and cation A in the block layer that separates two CuO_2 layers. (See Fig. 1 for an illustration of X and A.) The MOD up to the second order is[1]

$$|t_1|_{\rm MOD2} = 0.534 - 0.000327\left[R_{\rm X}^{9.99} + 7.99R_{\rm A}^{7.76}\right]. \quad (2)$$

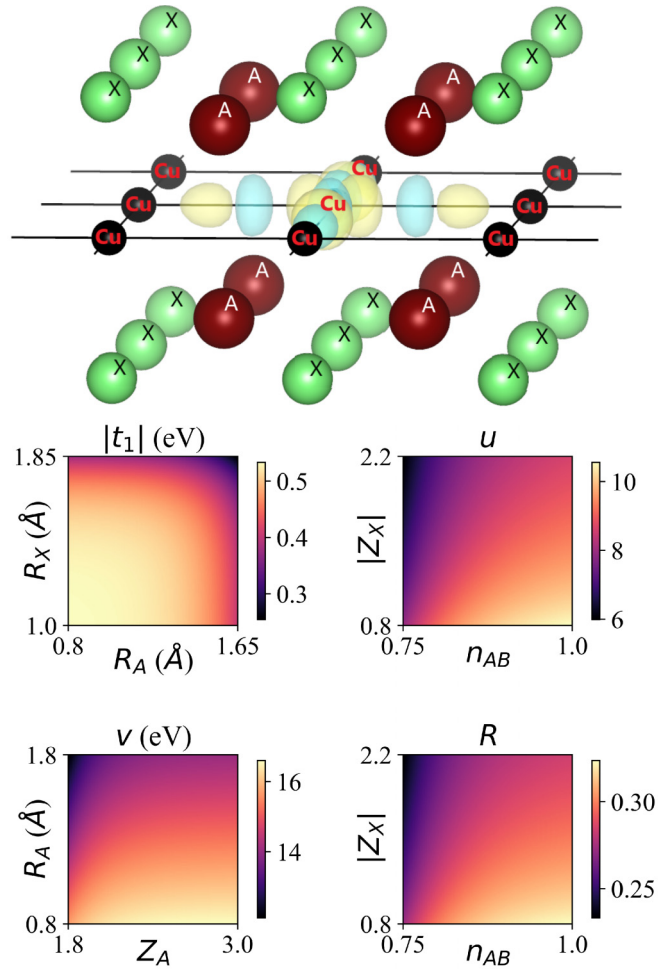

FIG. 1. (upper panel) Simplified representation of the CuO_2 plane and the surrounding crystalline environment. We show the square lattice formed by the Cu atoms in the CuO_2 plane (the in-plane O atoms are not shown), and the isosurface of the AB orbital centered on one of the Cu atoms (yellow is positive, blue is negative). We also show the A cations and apical X anions near the CuO_2 plane. (lower panel) Representation of the MODs of the AB Hamiltonian parameters on the CF obtained from gMACE + HDE. We show the MOD of $|t_1|$ [Eq. (2)] and $u = U/|t_1|$ [Eq. (3)] together with the onsite bare Coulomb interaction v [Eq. (4)] and the screening ratio $R = U/v$ [Eq. (5)]. $R_{\rm A}$ and $Z_{\rm A}$ ($R_{\rm X}$ and $Z_{\rm X}$) are the crystal ionic radius and ionic charge of the A cation (apical X anion). $n_{\rm AB}$ is the average number of electrons per AB orbital.

Qualitatively, $|t_1|$ increases when $R_{\rm X}$ or $R_{\rm A}$ decreases (see Fig. 1). The microscopic mechanism is summarized as follows: Reducing $R_{\rm X}$ and/or $R_{\rm A}$ reduces the chemical pressure that pushes atoms apart from each other inside the crystal. This reduces the cell parameter a and thus the distance between Cu atoms in the CuO_2 plane, which increases the overlap between AB orbitals located on neighboring Cu sites, and thus $|t_1|$. This result is consistent with that in Ref. [20], in which $|t_1|$ increases when applying physical uniaxial pressure P_a along a direction. (The application of P_a reduces a.)

(II) u mainly depends on the negative ionic charge $Z_{\rm X}$ of the apical anion and the average number of electrons $n_{\rm AB}$ in the AB orbital. (We have $n_{\rm AB} = 1 - \delta$, where δ is the hole

[1]In Eqs. (2), (3), (4), and (5), the unit of $|t_1|$ and v is eV, whereas u and R are dimensionless. The ranges of values obtained in our *ab initio* calculations are $|t_1| = 0.40$–0.57 eV, $u = 7.2$–10.6, $v = 13.3$–16.4 eV, and $R = 0.23$–0.34. These ranges of values are reproduced by Eqs. (2), (3), (4), and (5), as seen in Fig. 1.

doping.) The MOD up to the second order is

$$u_{\mathrm{MOD2}} = 200.725 - 190.294\left[|Z_{\mathrm{X}}|^{0.01} + 0.00155 n_{\mathrm{AB}}^{-7.99}\right]. \tag{3}$$

Qualitatively, u increases when $|Z_{\mathrm{X}}|$ decreases or n_{AB} increases (see Fig. 1). To understand the origin of the MOD of u in Eq. (3), we decompose $u = vR/|t_1|$, where v is the onsite bare Coulomb interaction and $R = U/v$ is the screening ratio, and we examine the MODs of v and R in (III) and (IV) below. (The MOD of u is mainly determined by the MOD of R.)

(III) v mainly depends on R_{A} and the positive ionic charge Z_{A} of the cation. The MOD up to the second order is

$$v_{\mathrm{MOD2}} = 18.874 - 2.787 R_{\mathrm{A}}^{0.92}\left[1 + 87.90 Z_{\mathrm{A}}^{-9.12}\right]. \tag{4}$$

Qualitatively, v increases when R_{A} decreases or Z_{A} increases (see Fig. 1). The microscopic mechanism is summarized as follows: Reducing R_{A} or increasing Z_{A} modifies the crystal electric field [namely, the Madelung potential (MP) created by cations and anions in the crystal] felt by the Cu3$d_{x^2-y^2}$ and O2p_σ electrons. This stabilizes the in-plane O2p_σ orbitals with respect to the Cu3$d_{x^2-y^2}$ orbitals, which increases the Cu3$d_{x^2-y^2}$/O2p_σ charge-transfer energy ΔE_{xp}. (Details are given later in Sec. IV.) This reduces the Cu3$d_{x^2-y^2}$/O2p_σ hybridization, which increases the Cu3$d_{x^2-y^2}$ atomic character of the AB orbital. This increases the localization of the AB orbital and thus v. This result is consistent with Ref. [20].

(IV) The CF dependence of R is more complex than that of $|t_1|$ and v, but we identify a rough MOD of R on Z_{X} and n_{AB}, which is

$$R_{\mathrm{MOD2}} = 4.224 - 3.906\left[|Z_{\mathrm{X}}|^{0.01} + 0.00078 n_{\mathrm{AB}}^{-9.99}\right]. \tag{5}$$

Qualitatively, R increases when (i) $|Z_{\mathrm{X}}|$ decreases or (ii) n_{AB} increases (see Fig. 1). The microscopic mechanism of (i) and (ii) is summarized below. (Details are given later in Sec. IV.)

(i) Decreasing $|Z_{\mathrm{X}}|$ reduces the negative charge of the apical anion. This reduces the negative MP created by the apical anion and felt by the electrons in the nearby CuO_2 plane. This reduces the energy of the electrons in the CuO_2 plane, and also reduces the Fermi energy. As a consequence, the empty states become higher in energy relative to the Fermi level. This reduces the screening from empty states, and thus, increases R.

(ii) The decrease in R with decreasing n_{AB} (increasing δ) is consistent with Ref. [19], in which the increase in δ causes the rapid decrease in R and thus u. (In Ref. [19], calculations were made at fixed $|Z_{\mathrm{X}}| = 2$ and varying $n_{\mathrm{AB}} = 1.0$, 0.9, and 0.8, which corresponds to $\delta = 0.0$, 0.1, and 0.2.) The rapid decrease in u eventually suppresses F_{SC} [14] so that the system ends up in the metallic state, in agreement with the experimental ground state in the overdoped region.

For completeness, the dependence of R on $|Z_{\mathrm{X}}|$ and n_{AB} beyond Eq. (5) is analyzed in Appendix G. Although R increases with decreasing $|Z_{\mathrm{X}}|$ in Eq. (5), note that a lower $|Z_{\mathrm{X}}|$ accelerates the decrease in R with decreasing n_{AB}. Due to this, R may decrease with decreasing $|Z_{\mathrm{X}}|$ in the overdoped region ($n_{\mathrm{AB}} \lesssim 0.85 - 0.80$), which is not captured by Eq. (5).

Remarkably, the MOD of R [Eq. (5)] is very similar to the MOD of u [Eq. (3)]. Thus, the MOD of $u = vR/|t_1|$ is dominated by the MOD of R. Consistently, in the *ab initio* result, the diverse distribution of $u \simeq 7$–10.5 originates mainly (albeit not exclusively) from the diverse distribution of R. Indeed, the relative variation between the minimum and maximum *ab initio* values is 27% for $v \simeq 13.5$–16.5 eV, 30% for $|t_1| \simeq 0.40$–0.55 eV, 48% for $R \simeq 0.22$–0.34, and 50% for $u \simeq 7$–10.5. (The variation in u is reproduced by that in R.) Still, the relative variation in $|t_1|$ and v is non-negligible compared with that in R, so that the CF dependencies of $|t_1|$ and v also contribute to the CF dependence of u beyond the MOD in Eq. (3). In Sec. IV, we decompose $u = vR/|t_1|$, and we discuss in detail the CF dependence of $|t_1|$, v, and R.

III. METHODOLOGY

A. Framework of hierarchical dependence extraction

Here, given a physical quantity y and other quantities x_i in the variable space $\mathcal{V} = \{x_i, i = 1, \ldots, N_{\mathcal{V}}\}$ (i is the variable index), we summarize the essence of the HDE procedure (denoted as HDE[y, $\mathcal{V}$]) to construct a descriptor for y as a function of x_i. (We make complete abstraction of the physical meaning of these variables.) Details on the motivation of the HDE and the HDE quantities and calculations are given in Appendix C.

The HDE expression of y is

$$y_N = k_0 + k_1 x_{(N)}^{\mathrm{opt}}, \tag{6}$$

where the HDE descriptor $x_{(N)}^{\mathrm{opt}}$ at generation $g = N$ is constructed iteratively by adding factors that contain dependencies of y on x_i from the lowest order to the highest order. At $g = 1$ and $g \geqslant 2$, we consider the candidate descriptors

$$x_{(1)} = x_{i_1}^{\alpha_1}, \tag{7}$$

$$x_{(g)} = x_{(g-1)}^{\mathrm{opt}} \star_{(\zeta_g, \beta_g, \alpha_g)} x_{i_g}, \tag{8}$$

where the parametric operator $\star_{(\zeta,\beta,\alpha)}$ (dubbed hereafter as the wildcard operator) is defined as

$$x \star_{(\zeta,\beta,\alpha)} x' = x\left[1 + \zeta \frac{x'^{\alpha}}{x^{\beta}}\right], \tag{9}$$

and $x_{(g)}^{\mathrm{opt}}$ is the best candidate descriptor $x_{(g)}$ at generation g. To determine $x_{(g)}^{\mathrm{opt}}$, we calculate the values $(i_g^{\mathrm{opt}}, \alpha_g^{\mathrm{opt}}, \zeta_g^{\mathrm{opt}}, \beta_g^{\mathrm{opt}})$ of the variational parameters $(i_g, \alpha_g, \zeta_g, \beta_g)$ that maximize the fitness function

$$f[y, x_{(g)}] = |\rho[y, x_{(g)}]|, \tag{10}$$

where ρ is the Pearson correlation coefficient. (The definition of ρ and further discussions on the choice of f are given in Appendix C 2, and computational details of the optimization are given in Appendix C 4.) At $g = 1$, we have only two variational parameters (i_1, α_1).

To obtain y_N in Eq. (6), we calculate $x_{(g)}^{\mathrm{opt}}$ starting from $g = 1$ and incrementing g up to N. We obtain

$$x_{(N)}^{\mathrm{opt}} = \left(x_{i_1^{\mathrm{opt}}}\right)^{\alpha_1^{\mathrm{opt}}} \prod_{g=2}^{N} \left[1 + \zeta_g^{\mathrm{opt}} \frac{\left(x_{i_g^{\mathrm{opt}}}\right)^{\alpha_g^{\mathrm{opt}}}}{\left[x_{(g-1)}^{\mathrm{opt}}\right]^{\beta_g^{\mathrm{opt}}}}\right], \tag{11}$$

and the coefficients k_0 and k_1 in Eq. (6) are calculated by an affine regression of y on $x_{(N)}^{\rm opt}$.

The wildcard operator in Eq. (9) is versatile and can represent any algebraic operation depending on the values of (ζ, β, α), as discussed in detail in Appendix C 3. In the practical procedure, $(\zeta_g, \beta_g, \alpha_g)$ are optimized together with i_g to maximize $f[y, x_{(g)}]$, so that the character of the wildcard operator is automatically adjusted to describe y as accurately as possible.

The HDE $[y, \mathcal{V}]$ allows us to probe the completeness of the dependence of y on x_i. When it is performed, we assume the following conjecture: $\mathcal{D}[y, \mathcal{V}]$ The dependence of y is entirely contained in $\mathcal{V} = \{x_i\}$. (This implies that there exists g and $x_{(g)}$ such that $f[y, x_{(g)}] = 1$.) The validity of $\mathcal{D}[y, \mathcal{V}]$ can be checked by examining the value of

$$f_\infty[y, \mathcal{V}] = \lim_{g\to\infty} f_{(g)}[y, \mathcal{V}], \tag{12}$$

in which

$$f_{(g)}[y, \mathcal{V}] = f\left[y, x_{(g)}^{\rm opt}\right]. \tag{13}$$

If $f_\infty[y, \mathcal{V}]$ is close to one, it is the proof that $\mathcal{D}[y, \mathcal{V}]$ is correct; in this case, the affine interpolation in Eq. (6) is accurate, and y_N is a good approximation of y at least for high values of N. If $f_\infty[y, \mathcal{V}]$ is not close to one, then either $\mathcal{D}[y, \mathcal{V}]$ is incorrect, or $\mathcal{D}[y, \mathcal{V}]$ is correct but the HDE $[y, \mathcal{V}]$ is insufficient to capture the whole dependence of y on the x_i. In the scope of this paper, we assume $\mathcal{D}[y, \mathcal{V}]$ is incorrect.

The expression of $x_{(N)}^{\rm opt}$ [Eq. (11)] reveals the hierarchy in the dependencies of y on x_i. Incrementing the generation index from $g-1$ to g introduces the variable $x_{i_g^{\rm opt}}$, which contains the gth-order dependence of y. Note that other variables may be in competition with $x_{i_g^{\rm opt}}$; such competition is examined by performing a score analysis in Appendix E.

In this paper, we mainly discuss the MODs that are contained in low values of $g \lesssim 2$–3. In this case, we write y_g in Eq. (6) as $y_{{\rm MOD}g}$, which is the MOD of y on x_i up to the gth order (MODg). The full list of variational parameters up to $g = 15$ is given in Sec. S1 of the Supplemental Material [21].

The accuracy of the MODg of y on x_i is quantified in $f_{(g)}[y, \mathcal{V}]$. In the general case, the MODg is but a rough description of y: Typically, $f_{(g)}[y, \mathcal{V}] \simeq 0.65$–$0.95$ for $g \lesssim 3$ (but sometimes $f_{(g)}[y, \mathcal{V}] \simeq 1$ as seen later). Accurate description requires to take into account the higher-order dependencies beyond $g \lesssim 3$ as well. Nonetheless, the MODg contains the principal mechanism of the dependence of y, as seen later.

In practice, the HDE$[y, \mathcal{V}]$ is employed as follows: We obtain $f_{(g)}[y, \mathcal{V}]$ and $(i_g^{\rm opt}, \alpha_g^{\rm opt}, \zeta_g^{\rm opt}, \beta_g^{\rm opt})$ for g from 1 to N. Then, we check the validity of $\mathcal{D}[y, \mathcal{V}]$ by examining the value of $f_\infty[y, \mathcal{V}]$. If $\mathcal{D}[y, \mathcal{V}]$ is incorrect, we may attempt to replace $\mathcal{V}$ by a superset of $\mathcal{V}$ and relaunch the procedure. If $\mathcal{D}[y, \mathcal{V}]$ is correct, we perform the below restricted procedure, denoted as rHDE$[y, \mathcal{V}]$. We attempt to simplify the dependence of y by eliminating the higher-order dependencies in decreasing order. Namely, we take the optimized variable indices $i_1^{\rm opt}, i_2^{\rm opt}, \ldots, i_N^{\rm opt}$ obtained at $g = 1, 2, \ldots, N$ in the HDE$[y, \mathcal{V}]$. By using these notations, HDE$[y, \mathcal{V}]$ is equivalent to HDE$[y, \{x_{i_1^{\rm opt}}, \ldots, x_{i_N^{\rm opt}}\}]$. Then, we perform HDE$[y, \{x_{i_1^{\rm opt}}, \ldots, x_{i_{N-j}^{\rm opt}}\}]$ by starting from $j = 1$ and incrementing j. (Each time we increment j, we remove the variable that corresponds to the highest-order dependence.) We check whether $\mathcal{D}_j = \mathcal{D}[y, \{x_{i_1^{\rm opt}}, \ldots, x_{i_{N-j}^{\rm opt}}\}]$ is correct. If $\mathcal{D}_j$ is correct but $\mathcal{D}_{j+1}$ is incorrect, we conclude that $\{x_{i_1^{\rm opt}}, \ldots, x_{i_{N-j}^{\rm opt}}\}$ is the minimal subset of $\mathcal{V}$ that describes y entirely.

B. Framework of gMACE

Next, we summarize the *ab initio* MACE scheme and its generalization to the gMACE in the present paper. [See Fig. 2 for an illustration.] Details are given in Appendix D.

Chemical formula dependence of crystal parameters. Obtaining the CF dependence of the AB Hamiltonian requires to extend the MACE scheme [15–20,22–24]. The latter allows us to calculate the AB Hamiltonian starting from a given CF together with the crystal symmetry and crystal parameter (CP) values. The CP values are usually taken from experiment, so that the missing step is to calculate *ab initio* the CP as a function of the CF. We add this step by introducing the CF variables (the radii and charges of cations and anions in the crystal) and calculating the CP as a function of the CF by performing the structural optimization, instead of relying on the experimental CP values. [Even though the experimental CP may be more accurate than the optimized CP, the experimental CP is not always available, and the structural optimization allows us to obtain the systematic CF dependence of the CP.] We only assume the symmetry of the primitive cell during the structural optimization (see Appendix D 1 for details and the atomic positions).

Calculation of the AB Hamiltonian. After obtaining the CP for a given CF as described above, the AB Hamiltonian is calculated by following the MACE procedure as in Ref. [20], which is summarized below in the successive steps (i)–(v). This procedure combines the generalized gradient approximation (GGA) [25] and the constrained random-phase approximation (cRPA) [26,27], and is denoted as GGA + cRPA.

(i) Starting from the CF and the CP values, we first perform a density functional theory (DFT) calculation. We obtain the DFT electronic structure at the GGA level.

(ii) From the GGA electronic structure, we compute the maximally localized Wannier orbitals [28,29] that span the medium-energy (M) space. The M space consists in the 17 bands with Cu3d-like, in-plane O2p-like, and apical X2p-like character near the Fermi level (see Fig. 2). The band with the highest energy in the M space contains most of the AB character, and the bands outside the M space form the high-energy space. The orbital centered on the atom l and with m-like orbital character is denoted as (l, m), where the atom index l, l' takes the values Cu, O, O′, X (O and O′ denote the two in-plane O atoms in the unit cell.) The orbital index m, m' takes the values x^2-y^2, $3z^2-r^2$, xy, and yz for the Cu3$d_{x^2-y^2}$, Cu3$d_{3z^2-r^2}$, Cu3d_{xy}, and Cu3d_{yz} orbitals (the Cu3d_{zx} and Cu3d_{yz} orbitals are equivalent), p_σ, p_π, and p_z for the in-plane O2p_σ, O2p_π, and O2p_z orbitals, and p_x, p_z for the apical X2p_x and X2p_z orbitals (the X2p_y and X2p_x orbitals are equivalent). Note that O2p_σ, O′2p_σ, O2p_π, and O′2p_π correspond respectively to O2p_x, O′2p_y, O2p_y, and O′2p_x.

(iii) From the M space, we compute the AB maximally localized Wannier orbital by using a procedure that is detailed in Appendix D. The AB orbital centered on the Cu atom

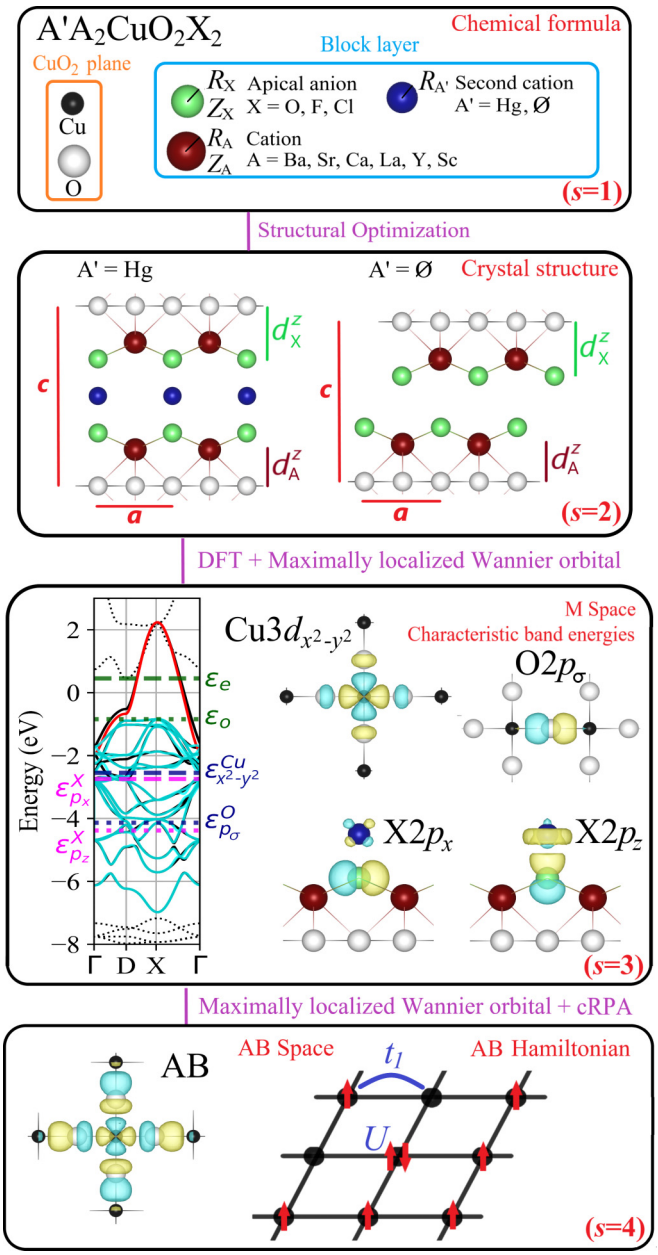

FIG. 2. Summary and illustration of the gMACE scheme and variables that are considered in the gMACE + HDE procedure. On $s = 2$, the primitive cell vectors and atomic positions are given in Appendix D 1. On $s = 3$, we show the band structure for $HgSr_2CuO_4$. [High-symmetry points are $\Gamma = (0, 0, 0)$, $D = (1/2, 0, 0)$, and $X = (1/2, 1/2, 0)$ in coordinates of the reciprocal lattice.] The dashed (solid) black bands are those outside (inside) the M space. The red band is the AB band, and the cyan bands are the 16 other M bands after disentanglement from the AB band. We also show the principal characteristic energies that are discussed in the main text: Energy ϵ_e (ϵ_o) of the lowest empty band (highest occupied disentangled M band) outside the AB band, and onsite energies $\epsilon^{Cu}_{x^2-y^2}$, $\epsilon^{O}_{p_\sigma}$, $\epsilon^{X}_{p_x}$, and $\epsilon^{X}_{p_z}$ of the maximally localized Wannier orbitals whose character is $Cu3d_{x^2-y^2}$, $O2p_\sigma$, $X2p_x$, and $X2p_z$. (The $Cu3d_{x^2-y^2}/O2p_\sigma$ charge-transfer energy is $\Delta E_{xp} = \epsilon^{Cu}_{x^2-y^2} - \epsilon^{O}_{p_\sigma}$). We also show the isosurfaces of these orbitals (yellow is positive, blue is negative). On $s = 4$, we show the isosurface of the AB orbital centered on the Cu atom, and a schematic illustration of the square lattice formed by the Cu atoms in the CuO_2 plane.

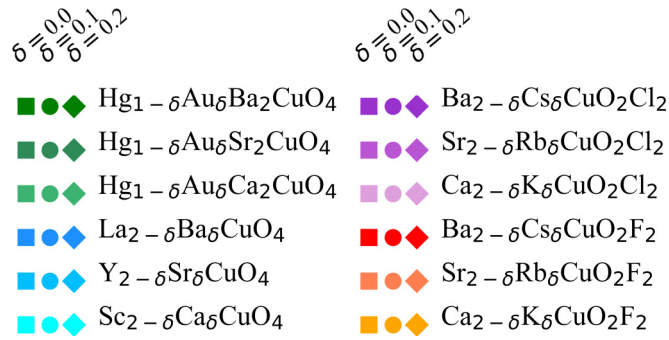

FIG. 3. List of the CFs in the training set of cuprates that is considered in this paper. We consider $N_t = 36$ CFs in total. The general CF is $A'A_2CuO_2X_2$. The color points that correspond to each CF are used in the subsequent figures.

located in the unit cell at **R** is denoted as $w_{\mathbf{R}}$, where **R** indicates the position of the unit cell in coordinates of the primitive vectors **a**, **b**, **c** of the Bravais lattice. We also obtain the AB band that corresponds to the dispersion of the AB orbital. Then, the other 16 bands in the M space are disentangled [30] from the AB band. (See Fig. 2 for an illustration of the AB band and disentangled M band dispersions.)

(iv) From the AB orbital, we compute $|t_1|$ as

$$t_1 = \int_\Omega dr w^*_{\mathbf{0}}(r) h^{\mathrm{GGA}}(r) w_{\mathbf{R}_1}(r), \tag{14}$$

in which Ω is the unit cell, $\mathbf{R}_1 = [100]$, and h^{GGA} is the one-particle part at the GGA level. We also compute v as

$$v = \int_\Omega dr \int_\Omega dr' w^*_{\mathbf{0}}(r) w^*_{\mathbf{0}}(r') v(r, r') w_{\mathbf{0}}(r) w_{\mathbf{0}}(r'), \tag{15}$$

in which $v(r, r')$ is the bare Coulomb interaction.

(v) From the AB band, the disentangled M bands and the high-energy bands outside the M space, we compute the cRPA screening ratio R and also u as follows. We compute the cRPA effective interaction $W_{\mathrm{H}}(r, r')$ at zero frequency, by using Eqs. (D1) and (D4) in Appendix D 1. We deduce U by replacing $v(r, r')$ with $W_{\mathrm{H}}(r, r')$ in Eq. (15). We deduce $R = U/v$ and $u = U/|t_1|$.

C. Training set of cuprates and gMACE + HDE procedure

Below, we define the training set considered in this paper. (For each CF in the training set, we apply the gMACE procedure described in Sec. III B.) Then, we describe how the HDE is applied to analyze the gMACE results and extract the CF dependence of the AB Hamiltonian parameters.

Training set. The training set is defined below and illustrated in Fig. 3. (Detailed discussions on the choice of the training set are given in Appendix B.) The training set includes $N_{\mathrm{tr}} = 36$ CFs, including both experimentally confirmed and hypothetical superconducting cuprates. The general CF is $A'A_2CuO_2X_2$ and the block layer consists in $A'A_2X_2$. For the undoped compound, we have A′ = Hg, ∅, A = Ba, Sr, Ca, La, Y, Sc, and X = O, F, Cl. For the doped compound, we use the same procedure as in Ref. [19,20]: We use the virtual crystal approximation [31] to substitute part of A or A′ by the chemical element whose atomic number is that of A or A′ minus one. We consider hole doping $\delta = 0.0$,

0.1, and 0.2 (that is, up to 20%). This range includes the experimental range in which the superconducting state is observed. We have $Z_{A'} = 2 - \delta$ if $A' = Hg_{1-\delta}Au_\delta$ and $Z_{A'} = 0$ if $A' = \varnothing$, $Z_A = 2 - \delta/2$ if A = Ba, Sr, Ca and $Z_A = 3 - \delta/2$ if A = La, Y, Sc, $Z_X = -2$ if X = O and $Z_X = -1$ if X = F, Cl. The ionic charges are related to δ and n_{AB} as follows:

$$Z_{A'} + 2Z_A + 2Z_X = 2 - \delta = 1 + n_{AB}. \tag{16}$$

Note that all CFs in the training set do not need to correspond to experimentally confirmed superconducting cuprates. For a given CF, the gMACE result can be used as part of the data that are analyzed by the HDE procedure to infer the systematic CF dependence of $|t_1|$ and u by making complete abstraction of whether the corresponding crystal structure can be stabilized in experiment.

gMACE + HDE procedure. We apply the HDE to express the AB Hamiltonian parameters as a function of the CF variables. In addition, to gain further insights on the underlying microscopic mechanism, we apply the HDE to express the intermediate quantities within the gMACE as functions of each other. Namely, we consider $N_s = 4$ levels of variables whose definition is guided by the hierarchical structure of gMACE as illustrated in Fig. 2. At each step s, we define the variable space $\mathcal{V}_s = \{x_i^s, i = 1 \ldots, N_{\mathcal{V}_s}\}$, and we use the HDE to express x_i^s as a function of the variables in $\mathcal{V}_{s'}$ (with $s' < s$). For each s, the variables in $\mathcal{V}_s$ are chosen as follows. (These variables are illustrated in Fig. 2.)

Chemical formula ($s = 1$). We consider $\mathcal{V}_1 = \{R_A, R_X, Z_A, |Z_X|, R_{A'}, n_{AB}\}$. $R_{A'}$ is the crystal ionic radius of the second cation $A' = Hg_{1-x}Au_x$ in the block layer ($R_{A'}$ is set to 0 if $A' = \varnothing$). Values of the variables in $\mathcal{V}_1$ considered in this paper are given in Appendix D 3. R_A, R_X, and $R_{A'}$ are expressed in Å in this paper.

Crystal parameters ($s = 2$). We consider $\mathcal{V}_2 = \{a, c, d_A^z, d_X^z, c_\perp\}$, where a, c, and $c_\perp$ are the cell parameters, and d_A^z (d_X^z) is the distance between the CuO_2 plane and the A cation (apical X anion). The coordinates of primitive vectors and atoms in the unit cell are given and discussed in Appendix D 1. All variables in $\mathcal{V}_2$ are expressed in Å in this paper.

DFT band structure ($s = 3$). We consider $\mathcal{V}_3 = \{|\epsilon_m^l|, |t_{m,m'}^{l,l'}|, \Delta E_{xp}, W_M, |\epsilon_o|, \epsilon_e\}$. These variables are characteristic energies within the M space and are defined below. [Their choice is further discussed in Appendix D 2.] First, we include the absolute values of the onsite energies ϵ_m^l of all maximally localized Wannier orbitals (l, m) in the M space. (Note that $\epsilon_m^l < 0$ because the onsite energy is below the Fermi level.) Second, we include the nonzero hopping amplitudes $|t_{m,m'}^{l,l'}|$ between the orbitals (l, m) and (l', m') in the unit cell, where $(l, m) = (Cu, x^2 - y^2)$, (O, p_σ). We exclude hoppings in the unit cell that do not involve $(Cu, x^2 - y^2)$ or (O, p_σ) orbitals, and hoppings beyond the unit cell. We use the abbreviation $|t_{xp}| = |t_{x^2-y^2,p_\sigma}^{Cu,O}|$. Third, we include the charge-transfer energy $\Delta E_{xp} = \epsilon_{x^2-y^2}^{Cu} - \epsilon_{p_\sigma}^{O}$ between the $(Cu, x^2 - y^2)$ and (O, p_σ) orbitals, W_M is the bandwidth of the M space, and ϵ_o (ϵ_e) is the energy of the highest occupied band in the M space (lowest empty band among the high-energy bands outside the M space) outside the AB band. (Note that $\epsilon_o < 0$ and $\epsilon_e > 0$.) All variables in $\mathcal{V}_3$ are expressed in eV in this paper.

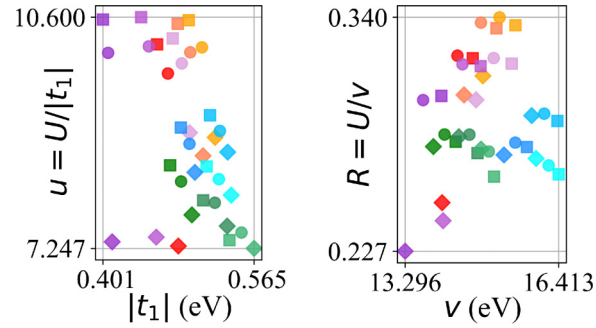

FIG. 4. *Ab initio* values of $|t_1|$, $u = U/|t_1|$, v, and $R = U/v$ obtained by employing the gMACE for all CFs in the training set. For each color point, the corresponding CF is shown in Fig. 3.

AB Hamiltonian parameters ($s = 4$). We consider $\mathcal{V}_4 = \{|t_1|, u, v, R\}$ as discussed in Sec. II. $|t_1|$ and v are expressed in eV in this paper.

IV. RESULTS

The *ab initio* values of $|t_1|$ and u together with those of v and R are summarized in Fig. 4. In this section, we detail the dependencies of $|t_1|$, v, and R on $\mathcal{V}_1$. These dependencies correspond to the items (I), (III), and (IV) discussed in Sec. II and shown in Fig. 5. We also detail the microscopic mechanism of (I), (III), and (IV) by detailing the dependencies between intermediate quantities within the gMACE [the items (1)–(10) in Fig. 5].

A. Chemical formula dependence of $|t_1|$

Here, we first detail (I). Then, we detail the items (1)–(5) in Fig. 5.

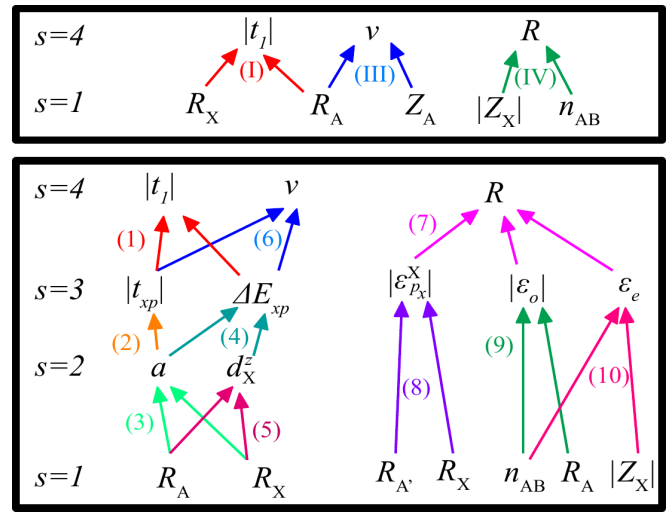

FIG. 5. Summary of the MODs between quantities obtained within gMACE + HDE. (upper panel) MOD2s of $|t_1|$, v, and R on the CF variables in $\mathcal{V}_1$. [The items (I), (III), and (IV) have been summarized in Sec. II, and are discussed in detail in Sec. IV.] (lower panel) MODs between intermediate quantities within the gMACE. [The items (1)–(10) are discussed in detail in Sec. IV.]

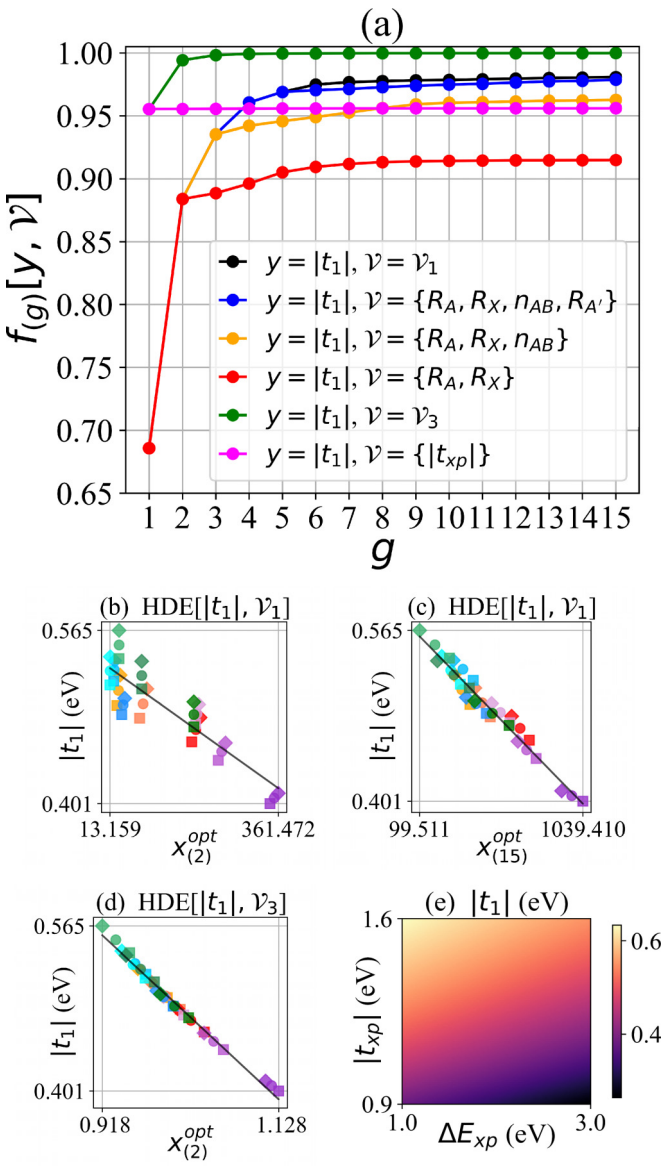

FIG. 6. Details on the dependence of $|t_1|$ on $\mathcal{V}_1$ and $\mathcal{V}_3$. (a) Values of $f_{(g)}[|t_1|, \mathcal{V}_1]$ in the HDE[$|t_1|, \mathcal{V}_1$], the rHDE[$|t_1|, \mathcal{V}_1$], the HDE[$|t_1|, \mathcal{V}_3$], and the rHDE[$|t_1|, \mathcal{V}_3$]. (b), (c) Dependence of $|t_1|$ on $x^{\rm opt}_{(2)}$ [corresponding to Eq. (2)] and $x^{\rm opt}_{(15)}$ in the HDE[$|t_1|, \mathcal{V}_1$]. (d) Dependence of $|t_1|$ on $x^{\rm opt}_{(2)}$ in Eq. (17). (e) Representation of the MOD2 of $|t_1|$ on $\{|t_{xp}|, \Delta E_{xp}\}$ in Eq. (17). In the panels (b)–(d), the CF that corresponds to each color point is shown in Fig. 3, and the solid black line shows the linear interpolation.

(I) *Dependence of* $|t_1|$ *on* $\mathcal{V}_1$. First, $|t_1|$ depends entirely on the variables in $\mathcal{V}_1$. The HDE[$|t_1|, \mathcal{V}_1$] yields $f_\infty[|t_1|, \mathcal{V}_1] = 0.981$ [see Fig. 6(a)], so that $\mathcal{D}[|t_1|, \mathcal{V}_1]$ is correct. Consistently, at $g = 15$, the dependence of $|t_1|$ on $x^{\rm opt}_{(15)}$ is almost linear [see Fig. 6(c)].

Second, the dependence of $|t_1|$ on $\mathcal{V}_1$ can be restricted to the subset $\{R_{\rm A}, R_{\rm X}, n_{\rm AB}, R_{\rm A'}\}$ of $\mathcal{V}_1$. In the HDE[$|t_1|, \mathcal{V}_1$], the variables $R_{\rm X}$, $R_{\rm A}$, $R_{\rm A'}$, and $n_{\rm AB}$ correspond to $x_{i_1^{\rm opt}}$, $x_{i_2^{\rm opt}}$, $x_{i_3^{\rm opt}}$, and $x_{i_4^{\rm opt}}$, respectively. The rHDE[$|t_1|, \mathcal{V}_1$] yields $f_\infty[|t_1|, \{R_{\rm A}, R_{\rm X}, R_{\rm A'}, n_{\rm AB}\}] = 0.979$ and $f_\infty[|t_1|, \{R_{\rm A}, R_{\rm X}, R_{\rm A'}\}] = 0.963$ [see Fig. 6(a)]. Thus, $\mathcal{D}[|t_1|, \{R_{\rm A}, R_{\rm X}, R_{\rm A'}, n_{\rm AB}\}]$ is correct, but $\mathcal{D}[|t_1|, \{R_{\rm A}, R_{\rm X}, R_{\rm A'}\}]$ is not. Thus, $\{R_{\rm A}, R_{\rm X}, R_{\rm A'}, n_{\rm AB}\}$ is the minimal subset of $\mathcal{V}_1$ that describes $|t_1|$.

Third, we discuss details of the MOD2 of $|t_1|$ on $\mathcal{V}_1$, which was given in Eq. (2). The MOD2 is not sufficient to describe $|t_1|$ entirely, but the main-order dependence of $|t_1|$ is captured. We have $f_{(2)}[|t_1|, \mathcal{V}_1] = 0.884$ [see Fig. 6(a)], and the dependence of $|t_1|$ on $x^{\rm opt}_{(2)}$ is shown in Fig. 6(b). Note that the dependence of $|t_1|$ on $R_{\rm A}$ is equally important to that on $R_{\rm X}$ in Eq. (2), even though $R_{\rm X}$ corresponds to $x_{i_1^{\rm opt}}$ (see the score analysis in Appendix E). In Appendix F, the robustness of Eq. (2) and the case in which we assume $(x_{i_1^{\rm opt}}, x_{i_2^{\rm opt}}) = (R_{\rm A}, R_{\rm X})$ instead of $(R_{\rm X}, R_{\rm A})$ are further analyzed.

(1) Dependence of $|t_1|$ *on* $\mathcal{V}_3$. First, $|t_1|$ is entirely determined by $|t_{xp}|$ and ΔE_{xp} irrespective of other variables in $\mathcal{V}_3$. The HDE[$|t_1|, \mathcal{V}_3$] yields $f_\infty[|t_1|, \mathcal{V}_3] = 1.000$ [see Fig. 6(a)]. The variables $|t_{xp}|$ and ΔE_{xp} correspond to $x_{i_1^{\rm opt}}$ and $x_{i_2^{\rm opt}}$, respectively. The rHDE[$|t_1|, \mathcal{V}_3$] yields $f_\infty[|t_1|, \{|t_{xp}|, \Delta E_{xp}\}] = 0.999$ and $f_\infty[|t_1|, \{|t_{xp}|\}] = 0.956$, so that $\mathcal{D}[|t_1|, \{|t_{xp}|, \Delta E_{xp}\}]$ is correct but $\mathcal{D}[|t_1|, \{|t_{xp}|\}]$ is not. Thus, $\{|t_{xp}|, \Delta E_{xp}\}$ is the subset of $\mathcal{V}_3$ that describes $|t_1|$.

Second, the MOD2 of $|t_1|$ on $\{|t_{xp}|, \Delta E_{xp}\}$ is sufficient to describe accurately $|t_1|$. Indeed, we have $f_{(2)}[|t_1|, \{|t_{xp}|, \Delta E_{xp}\}] = 0.994$, and the dependence of $|t_1|$ on $x^{\rm opt}_{(2)}$ is almost linear [see Fig. 6 (d)]. The MOD2 is

$$|t_1|_{\rm MOD2} = 1.269 - 0.777\left\{|t_{xp}|^{-0.58}\left[1 + 0.0739\Delta E_{xp}^{0.98}\right]\right\} \tag{17}$$

and is represented in Figs. 6(d) and 6(e). Note that $|t_{xp}|$ dominates over ΔE_{xp} in the MOD2 (see the score analysis in Appendix E), which is also visible in Fig. 6(e): The color map has a horizontal-like pattern.

Qualitatively, $|t_1|_{\rm MOD2}$ increases with increasing $|t_{xp}|$ and decreasing ΔE_{xp} in Eq. (17), and this is consistent with previous works. In the case of Hg1223, we have $|t_1| \propto |t_{xp}|^2/\Delta E_{xp}$ in Fig. 12 of Ref. [20]. (The latter result was obtained by modifying a artificially without modifying other CPs.) The result in Eq. (17) is more general, because Eq. (17) is established for $N_{\rm tr} = 36$ CFs rather than one CF, and it accounts for the CF dependence of all CPs.

The MOD2 in Eq. (17) can be interpreted as follows. $|t_1|$ represents the hopping amplitude between neighboring AB orbitals, and thus, $|t_1|$ mainly depends on the overlap between neighboring AB orbitals. Since the AB orbital is formed by the Cu3$d_{x^2-y^2}$ and in-plane O2p_σ orbitals, the overlap between the neighboring AB orbitals is determined by the overlap between the neighboring Cu3$d_{x^2-y^2}$ and in-plane O2p_σ orbitals, which is mainly encoded in $|t_{xp}|$. Thus, it is natural that $|t_1|$ mainly depends on $|t_{xp}|$. In addition, decreasing ΔE_{xp} reduces the localization of the AB orbital (as discussed later in Sec. IV B): The delocalization of the AB orbital within the CuO_2 plane contributes to increase the overlap between neighboring AB orbitals and thus $|t_1|$.

(2) Dependence of $|t_{xp}|$ *on* $\mathcal{V}_2$. $|t_{xp}|$ is entirely determined by the cell parameter a irrespective of other variables in $\mathcal{V}_2$, and the MOD1 describes $|t_{xp}|$ perfectly. Indeed, the HDE[$|t_{xp}|, \mathcal{V}_2$] yields $f_{(1)}[|t_{xp}|, \mathcal{V}_2] = 1.000$ [see Fig. 7(a)],

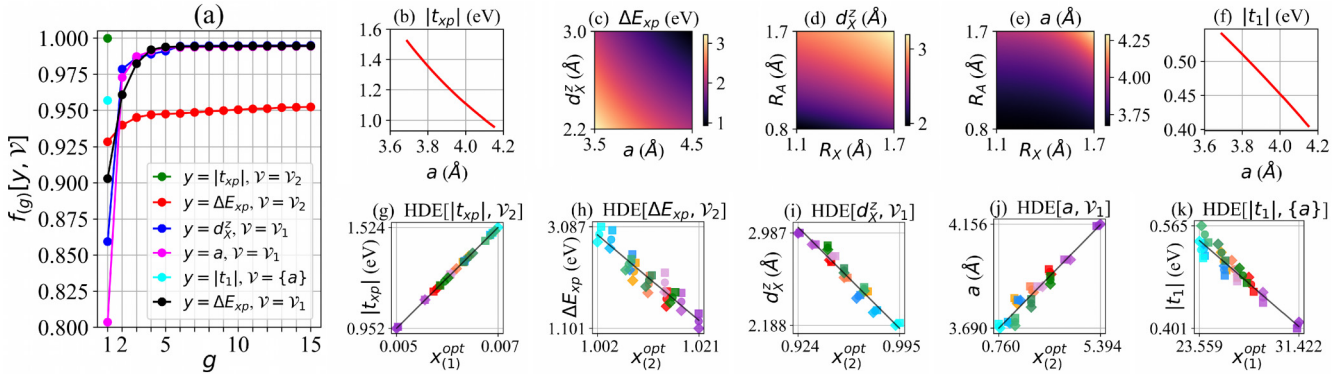

FIG. 7. Details on the dependence of $|t_{xp}|$ on $\mathcal{V}_2$, ΔE_{xp} on $\mathcal{V}_2$, a on $\mathcal{V}_1$, d^z_X on $\mathcal{V}_1$, and also $|t_1|$ on $\mathcal{V}_1$ and ΔE_{xp} on $\mathcal{V}_1$ for completeness. (a) Values of $f_{(g)}[y, \mathcal{V}]$ in the HDE[$|t_{xp}|$, $\mathcal{V}_2$], the HDE[ΔE_{xp}, $\mathcal{V}_2$], the HDE[a, $\mathcal{V}_1$], the HDE[d^z_X, $\mathcal{V}_1$], the HDE[$|t_1|$, {a}], and the HDE[ΔE_{xp}, $\mathcal{V}_1$]. (b)–(f) Representation of the MOD1 of $|t_{xp}|$ on $\mathcal{V}_2$ [Eq. (18)], the MOD2 of ΔE_{xp} on $\mathcal{V}_2$ [Eq. (21)], the MOD2 of d^z_X on $\mathcal{V}_1$ [Eq. (22)], the MOD2 of a on $\mathcal{V}_1$ [Eq. (20)], and the MOD1 of $|t_1|$ on {a} [Eq. (19)]. (g)–(k) Dependence of $|t_{xp}|$ and $|t_1|$ on their respective $x^{\rm opt}_{(1)}$, and ΔE_{xp}, d^z_X, and a on their respective $x^{\rm opt}_{(2)}$. [Panels (g) to (k) correspond to panels (b) to (f), respectively.] The CF that corresponds to each color point is shown in Fig. 3, and the solid black line shows the linear interpolation.

and $x_{t_1^{\rm opt}}$ corresponds to a. The MOD1 is

$$|t_{xp}|_{\rm MOD1} = -0.0632 + 212.234a^{-3.75} \tag{18}$$

and is represented in Figs. 7(b) and 7(g). The score analysis shows that the dominance of a in the dependence of $|t_{xp}|$ is unambiguous (see Appendix E).

The a dependence of $|t_{xp}|$ is consistent with results on Hg1223 [20]. In particular, the exponent -3.75 in Eq. (18) is very close to that obtained for Hg1223, in which $|t_{xp}| \propto a^{-3.86}$ (see Ref. [20], Fig. 12). This suggests $|t_{xp}|$ scales as $a^{-3.75}$ universally and irrespective of the crystalline environment outside the CuO_2 plane. This is intuitive because the Cu3$d_{x^2-y^2}$ and O2p_σ orbitals extend mainly in the CuO_2 plane as illustrated in Fig. 2.

In addition, $|t_1|$ increases when a decreases according to Eqs. (18) and (17). This is consistent with Ref. [20], in which the pressure-induced decrease in a is the main cause of the pressure-induced increase in $|t_1|$. For completeness, we perform the HDE[$|t_1|$, {a}]: We obtain $f_\infty[|t_1|, \{a\}] = f_{(1)}[|t_1|, \{a\}] = 0.957$, so that a describes $|t_1|$ reasonably, but not perfectly. This is because $|t_1|$ has not only a dominant dependence on $|t_{xp}|$ but also a small dependence on ΔE_{xp}, and ΔE_{xp} is not described entirely by a as seen later in (4). The MOD1 of $|t_1|$ on a is

$$|t_1|_{\rm MOD1} = 0.956 - 0.0176a^{2.42}, \tag{19}$$

and is represented in Figs. 7(f) and 7(k). Qualitatively, $|t_1|_{\rm MOD1}$ increases when a decreases, which is consistent with Ref. [20] and also with Eqs. (18) and (17). For completeness, note that there is a quantitative difference between Eq. (19) and [20]: In the latter, we have $|t_1| \propto a^{-2.88}$ (see Ref. [20], Fig. 12). The difference may be explained as follows. $|t_1|$ depends on both $|t_{xp}|$ and ΔE_{xp} [Eq. (17)], and ΔE_{xp} depends on the crystalline environment outside the CuO_2 plane contrary to $|t_{xp}|$. [For instance, as seen later in (5), the MOD2 of ΔE_{xp} depends not only on a but also on d^z_X.] The result in Ref. [20] captures the a dependence of $|t_1|$ by fixing the other CP values; the present result is more general because it accounts for the materials dependence of other CP values *via* the structural optimization.

(3) Dependence of a on $\mathcal{V}_1$. a is determined entirely by R_A and R_X in $\mathcal{V}_1$. The HDE[a, $\mathcal{V}_1$] yields $f_\infty[a, \mathcal{V}_1] = 0.994$, and the rHDE[$a$, $\mathcal{V}_1$] yields $f_\infty[a, \{R_A, R_X\}] = 0.985$. On the MOD2, we have $f_{(2)}[a, \{R_A, R_X\}] = 0.973$, and

$$a_{\rm MOD2} = 3.613 + 0.100\left\{R_A^{2.71}\left[1 + 0.00711R_X^{9.14}\right]\right\}, \tag{20}$$

which is represented in Fig. 7(e) and 7(j). The score analysis confirms that R_A and R_X correspond respectively to $x_{t_1^{\rm opt}}$ and $x_{t_2^{\rm opt}}$ (see Appendix E).

The MOD2 in Eq. (20) is interpreted as follows: Qualitatively, $a_{\rm MOD2}$ increases when R_A or R_X increases, which is consistent with the hard-sphere picture illustrated in Fig. 8: The interatomic distances and thus the cell parameter a increase when the ionic radii are larger. Thus, the hard-sphere picture reproduces the qualitative dependence of a on the ionic radii. Furthermore, the MOD2 in Eq. (20) is qualitatively consistent with experiment, in which a increases with increasing R_A. For instance, for X = Cl, we have $a = 3.87$ Å for A = Ca [32,33], $a = 3.97$–3.98 Å for A = Sr [34,35], and $a = 4.10$ Å for A = Ba [10]. (These values are in correct agreement with the values $a = 3.88$ Å for A = Ca, $a = 4.00$ Å for A = Sr and $a = 4.15$ Å for A = Ba obtained in this

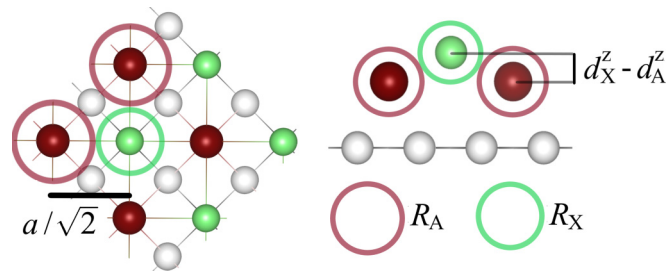

FIG. 8. Illustration of the hard-sphere picture for the crystal of the single-layer cuprate. In this picture, cations, and apical anions are assumed to be rigid spheres that touch each other, and we have $R_A + R_X = [(a/\sqrt{2})^2 + (d^z_X - d^z_A)^2]^{1/2}$.

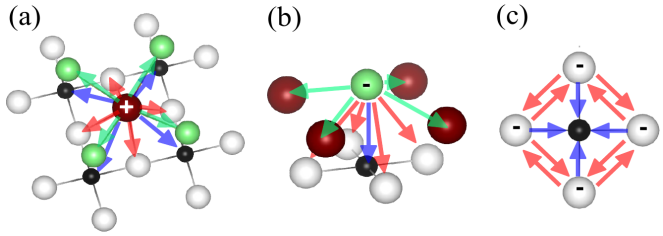

FIG. 9. Schematic illustration of the positive Madelung potential (MP) and negative MP created by the cations and anions within the crystal. (a) Positive MP created by the cation A and felt by the nearest-neighbor apical X (green arrows), in-plane O (red arrows), and Cu (blue arrows). (b) negative MP created by the apical X anion and felt by the nearest-neighbor Cu (blue arrow), in-plane O (red arrows), and cations (green arrows). (c) negative MP created by the in-plane O anion and felt by the nearest-neighbor Cu (blue arrows) and other in-plane O (red arrows). In panel (a), the positive MP created by A and felt at a distance d of A scales as $V_A(d) = Z_A/d$. The distance $d_{A,X} = [a^2/2 + (d_X^z - d_A^z)^2]^{1/2}$ between A and X is smaller than the distance $d_{A,O} = [a^2/4 + (d_A^z)^2]^{1/2}]$ between A and in-plane O, which is smaller than the distance $d_{A,Cu} = [a^2/2 + (d_A^z)^2]^{1/2}$ between A and Cu. Thus, we have $V_A(d_{A,X}) > V_A(d_{A,O}) > V_A(d_{A,Cu})$. In the panel (b), we have $|V_X(d_{X,A})| > |V_X(d_{X,Cu})| > |V_X(d_{X,O})|$. In the panel (c), we have $|V_O(d_{O,Cu})| > |V_O(d_{O,O'})|$.

paper by employing the structural optimization.) And we have $R_{Ca} = 1.14$ Å, $R_{Sr} = 1.32$ Å and $R_{Ba} = 1.49$ Å [36]. (The values of R_A are given in Appendix D 3, and the R_A dependence of experimental a is also emphasized in Ref. [10].) Also, for X = O, we have $a \simeq 3.78$ Å in La_2CuO_4 [37] and $a \simeq 3.88$ Å in $HgBa_2CuO_4$ [4]. (These values are in correct agreement with the values $a \simeq 3.82$ in La_2CuO_4 and $a \simeq 3.92$ Å in $HgBa_2CuO_4$ obtained in this paper by employing the structural optimization.) And, we have $R_{La} = 1.17$ Å and $R_{Ba} = 1.49$ Å [36].

(4) Dependence of ΔE_{xp} on $\mathcal{V}_2$. ΔE_{xp} is not determined entirely by $\mathcal{V}_2$, but the MOD2 reveals the main-order mechanism that controls the value of ΔE_{xp}. The HDE$[\Delta E_{xp}, \mathcal{V}_2]$ yields $f_\infty[\Delta E_{xp}, \mathcal{V}_2] = 0.952$ [see Fig. 7(a)], so that $\mathcal{D}[\Delta E_{xp}, \mathcal{V}_2]$ is not completely correct. As for the MOD2, we have $f_{(2)}[\Delta E_{xp}, \mathcal{V}_2] = 0.940$; the MOD2 is

$$\Delta E_{xp\mathrm{MOD2}} = 88.343 - 85.265\left[\left(d_X^z\right)^{0.04} - 0.479a^{-2.12}\right] \tag{21}$$

and is represented in Figs. 7(c) and 7(h). Even though d_X^z corresponds to $x_{i_1^{\mathrm{opt}}}$, the dependence of ΔE_{xp} on d_X^z is equally important to that on a (see the score analysis in Appendix E). Consistently, in Fig. 7(c), the color map has a diagonal-like pattern.

Qualitatively, ΔE_{xp} increases when d_X^z decreases or a decreases in Eq. (21). The increase in ΔE_{xp} with decreasing d_X^z can be understood as follows. When d_X^z decreases, the distance between the apical X and the CuO_2 plane decreases, so that the negative MP created by the apical X anion and felt by the Cu and in-plane O is stronger. This increases the energy of both Cu3$d_{x^2-y^2}$ and O2p_σ electrons. The negative MP felt by Cu is the strongest, because the Cu is closer to the apical X compared with the in-plane O. [See Fig. 9(b) for an illustration.] Thus, the energy of Cu3$d_{x^2-y^2}$ electrons increases more than that of in-plane O2p_σ electrons. As a consequence, ΔE_{xp} increases.

The increase in ΔE_{xp} with decreasing a is consistent with Ref. [20], and the mechanism is reminded here. When a decreases, the distance between the in-plane O and Cu is reduced, so that the negative MP created by the in-plane O anions and felt by the Cu is stronger. [See Fig. 9(c) for an illustration.] This increases the energy of Cu3$d_{x^2-y^2}$ electrons with respect to that of in-plane O2p_σ electrons, which increases ΔE_{xp}.

(5) Dependence of d_X^z on $\mathcal{V}_1$. d_X^z is determined entirely by $\mathcal{V}_1$. The HDE$[d_X^z, \mathcal{V}_1]$ yields $f_\infty[d_X^z, \mathcal{V}_1] = 0.995$. As for the MOD2 of d_X^z on $\mathcal{V}_1$, we have $f_{(2)}[d_X^z, \mathcal{V}_1] = 0.979$, and

$$d_{X\mathrm{MOD2}}^z = 14.353 - 12.254\left\{R_A^{-0.10}\left[1 - 0.00897R_X^{2.80}\right]\right\} \tag{22}$$

is represented in Figs. 7(d) and 7(i). The score analysis confirms that R_A and R_X correspond respectively to $x_{i_1^{\mathrm{opt}}}$ and $x_{i_2^{\mathrm{opt}}}$ (see Appendix E).

Qualitatively, $d_{X\mathrm{MOD2}}^z$ increases when R_A or R_X increases in Eq. (22). This can also be understood by considering the hard-sphere picture (see the right panel in Fig. 8). In the *ab initio* result, we always have $d_A^z < d_X^z$ (the values of d_A^z and d_X^z are given in Appendix D 3), so that the apical X is farther from the CuO_2 plane compared with the A cation. In the hard-sphere picture, increasing the ionic radius R_A of the A cation pushes the apical X even farther from the CuO_2 plane, which increases d_X^z. The same mechanism occurs when R_X increases.

Dependence of ΔE_{xp} on $\mathcal{V}_1$. For (4) and (5), for completeness, we discuss the dependence of ΔE_{xp} on $\mathcal{V}_1$. ΔE_{xp} is determined entirely by the subset $\{R_A, R_X, n_{AB}\}$ of $\mathcal{V}_1$. The HDE$[\Delta E_{xp}, \mathcal{V}_1]$ yields $f_\infty[\Delta E_{xp}, \mathcal{V}_1] = 0.995$, and the rHDE$[\Delta E_{xp}, \mathcal{V}_1]$ yields $f_\infty[\Delta E_{xp}, \{R_A, R_X, n_{AB}\}] = 0.988$ and $f_\infty[\Delta E_{xp}, \{R_A, R_X\}] = 0.967$ [see Fig. 7(a)]. Thus, $\mathcal{D}[\Delta E_{xp}, \{R_A, R_X, n_{AB}\}]$ is correct, but $\mathcal{D}[\Delta E_{xp}, \{R_A, R_X\}]$ is not. We have $f_{(2)}[\Delta E_{xp}, \mathcal{V}_1] = 0.961$ and $f_{(3)}[\Delta E_{xp}, \mathcal{V}_1] = 0.982$, and the MOD3 is

$$\Delta E_{xp\mathrm{MOD3}} = 9.710 - 7.565 \times \left\{R_A^{0.33}\left[1 + 0.000411R_X^{9.33}\right] - 0.0999n_{AB}^{1.77}\right\}. \tag{23}$$

The score analysis confirms that R_A and R_X correspond respectively to $x_{i_1^{\mathrm{opt}}}$ and $x_{i_2^{\mathrm{opt}}}$, and n_{AB} corresponds to $x_{i_3^{\mathrm{opt}}}$ but is slightly in competition with Z_A (see Appendix E).

Qualitatively, $\Delta E_{xp\mathrm{MOD3}}$ increases when R_A decreases, R_X decreases or n_{AB} increases in Eq. (23). This is consistent with Eqs. (21), (22), and (20). Also, ΔE_{xp} decreases when n_{AB} decreases (the hole doping δ increases). This is consistent with Ref. [19] and explained as follows. When δ increases, the hole doping of O sites increases. (The holes localize on O sites to form the Zhang-Rice singlet.) This reduces the negative charge of in-plane O anions. This reduces the negative MP created by the in-plane O and felt by the nearby Cu. [See Fig. 9(c) for an illustration.] This reduces the energy of Cu3d electrons, and also reduces the Fermi energy. (Indeed, the AB band at the Fermi level has Cu3$d_{x^2-y^2}$ character.) On the other

hand, the O2p electrons are less affected. Thus, the energy difference $\Delta E_{xp} = \epsilon^{\mathrm{Cu}}_{x^2-y^2} - \epsilon^{\mathrm{O}}_{p_\sigma}$ increases.

B. Chemical formula dependence of v

Here, we first detail (III). Then, we detail the item (6) in Fig. 5. [Items (2)–(5) have already been discussed in the previous section].

(III) *Dependence of v on $\mathcal{V}_1$.* First, v depends entirely on the variables in $\mathcal{V}_1$. The HDE$[v, \mathcal{V}_1]$ yields $f_\infty[v, \mathcal{V}_1] = 0.992$ [see Fig. 10(a)], so that $\mathcal{D}[v, \mathcal{V}_1]$ is correct. Second, the dependence of v can be restricted to the subset $\{R_\mathrm{A}, Z_\mathrm{A}, R_\mathrm{X}\}$ of $\mathcal{V}_1$. The variables R_A, Z_A, R_X, n_AB correspond to $x^{\mathrm{opt}}_{i_g}$ at $g = 1, 2, 3, 4$, respectively, and the rHDE[v, $\mathcal{V}_1$] yields $f_\infty[v, \{R_\mathrm{A}, Z_\mathrm{A}, R_\mathrm{X}, n_\mathrm{AB}\}] = 0.991$, $f_\infty[v, \{R_\mathrm{A}, Z_\mathrm{A}, R_\mathrm{X}\}] = 0.980$, and $f_\infty[v, \{R_\mathrm{A}, Z_\mathrm{A}\}] = 0.950$, so that $\mathcal{D}[v, \{R_\mathrm{A}, Z_\mathrm{A}, R_\mathrm{X}\}]$ is correct but $\mathcal{D}[v, \{R_\mathrm{A}, Z_\mathrm{A}\}]$ is not. The MOD2 of v on $\mathcal{V}_1$ is given in Eq. (4). The score analysis shows that R_A and Z_A correspond to $x^{\mathrm{opt}}_{i_1}$ and $x^{\mathrm{opt}}_{i_2}$ unambiguously (see Appendix E). In Appendix F, the robustness of Eq. (4) is confirmed.

(6) Dependence of v on $\mathcal{V}_3$. v is entirely determined by $|t_{xp}|$ and ΔE_{xp} in $\mathcal{V}_3$, but the clarification of this dependence is a bit more subtle and we discuss it in detail here. The HDE$[v, \mathcal{V}_3]$ yields $f_\infty[v, \mathcal{V}_3] = 0.999$ [see Fig. 10(a)]. The variable $|\epsilon^{\mathrm{Cu}}_{p_z}|$ corresponds to $x^{\mathrm{opt}}_{i_1}$, and we have $f_{(1)}[v, \mathcal{V}_3] = 0.976$. However, the physical MOD1 of v is not on $|\epsilon^{\mathrm{Cu}}_{p_z}|$ but rather on ΔE_{xp}, as discussed below. The score analysis in Appendix E shows that the three variables $|\epsilon^{\mathrm{O}}_{p_\sigma}|$, $|\epsilon^{\mathrm{O}}_{p_\pi}|$, and $\Delta E_{xp} = \epsilon^{\mathrm{Cu}}_{x^2-y^2} - \epsilon^{\mathrm{O}}_{p_\sigma}$ are in very close competition with $|\epsilon^{\mathrm{Cu}}_{p_z}|$ at $g = 1$. These four variables have a common point: They are all related to the onsite energies of in-plane O2p orbitals. The information that can be extracted from the above result is the following: v is primarily controlled by the energy of the in-plane O2p orbitals. We pinpoint the physical dependence as that on ΔE_{xp} by considering the result in Ref. [20]: v mainly depends on $R_{xp} = |t_{xp}|/\Delta E_{xp}$, and v increases when R_{xp} decreases (that is, when $|t_{xp}|$ decreases or ΔE_{xp} increases). The interpretation is reminded here: Decreasing R_{xp} reduces the Cu3$d_{x^2-y^2}$/O2p_σ hybridization, which increases the localization of the AB orbital and thus v.[2] To confirm the consistency with [20], we perform the HDE$[v, \{\Delta E_{xp}, |t_{xp}|\}]$. We obtain $f_{(2)}[v, \{\Delta E_{xp}, |t_{xp}|\}] = 0.985$, so that $\mathcal{D}[v, \{\Delta E_{xp}, |t_{xp}|\}]$ is correct and the MOD2 of v on $\{\Delta E_{xp}, |t_{xp}|\}$ is accurate. The MOD2 is

$$v_{\mathrm{MOD2}} = 10.119 + 4.710\left\{\Delta E_{xp}^{0.93}\left[1 - 0.399\left|t_{xp}^{0.69}\right|\right]\right\} \quad (24)$$

and is illustrated in Fig. 10(d). We choose to keep Eq. (24) as the final result for the dependence of v on $\mathcal{V}_3$. For completeness, we also perform the HDE$[v, \{R_{xp}\}]$. We obtain $f_\infty[v, \{R_{xp}\}] = 0.972$ and $f_{(1)}[v, \{R_{xp}\}] = 0.971$, and the MOD1 is

$$v_{\mathrm{MOD1}} = 6.824 + 5.941 R_{xp}^{-0.67}. \quad (25)$$

[2]If $|t_{xp}| \to 0$, the Cu3$d_{x^2-y^2}$ and O2p_σ orbitals do not overlap and thus do not hybridize. If $\Delta E_{xp} \to +\infty$, the difference between the Cu3$d_{x^2-y^2}$ and O2p_σ energy levels becomes very large, so that the hybridization becomes negligible.

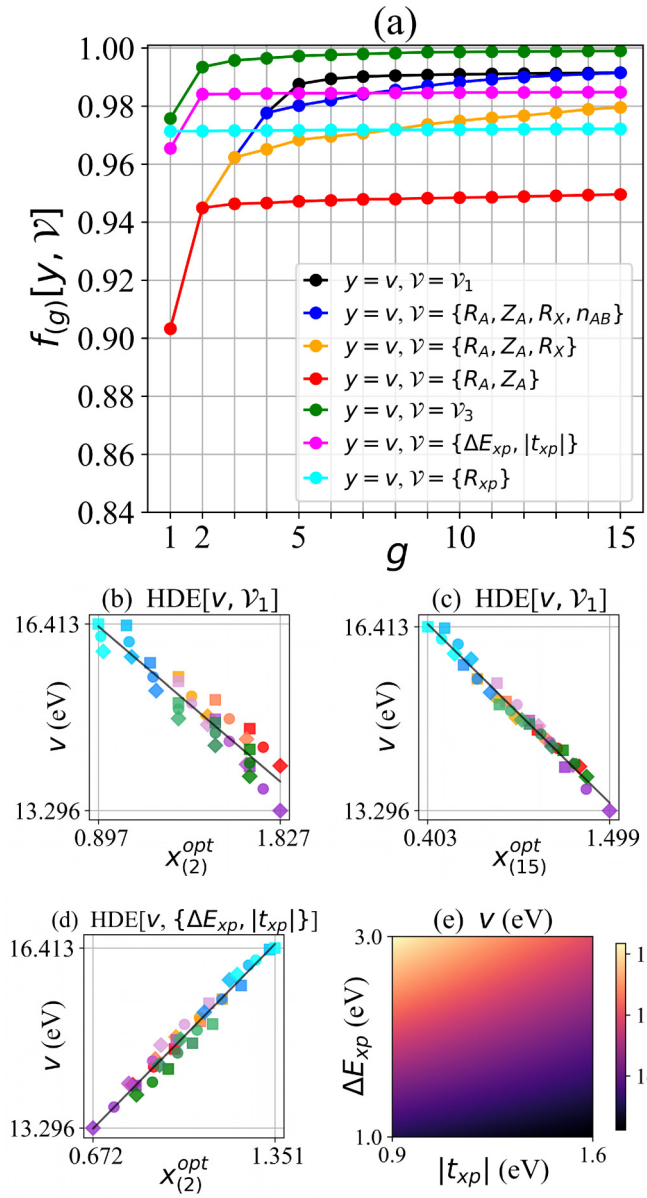

FIG. 10. Dependence of v on $\mathcal{V}_1$ and $\mathcal{V}_3$. (a) Values of $f_{(g)}[v, \mathcal{V}]$ in the HDE$[v, \mathcal{V}_1]$, the rHDE$[v, \mathcal{V}_1]$, the HDE$[v, \mathcal{V}_3]$, the HDE$[v, \{\Delta E_{xp}, |t_{xp}|\}]$ and the HDE$[v, \{R_{xp} = \Delta E_{xp}/|t_{xp}|\}]$. (b), (c) Dependence of $|t_1|$ on $x^{\mathrm{opt}}_{(2)}$ [corresponding to Eq. (4)] and $x^{\mathrm{opt}}_{(15)}$ in the HDE$[v, \mathcal{V}_1]$. (d) Dependence of $|t_1|$ on $x^{\mathrm{opt}}_{(2)}$ in the HDE$[v, \{\Delta E_{xp}, |t_{xp}|\}]$. Panel (d) corresponds to the MOD2 in Eq. (24). (e) Representation of the MOD2 of v on ΔE_{xp} and $|t_{xp}|$. In panels (b)–(d), the CF that corresponds to each color point is shown in Fig. 3, and the solid black line shows the linear interpolation.

According to both Eq. (24) and Eq. (25), v increases when $|t_{xp}|$ decreases or ΔE_{xp} increases, which is consistent with Ref. [20].

C. Chemical formula dependence of R

Here, we first detail (IV). Then, we discuss the items (7)–(10) in Fig. 5.

(IV) *Dependence of R on $\mathcal{V}_1$.* The dependence of R is more complex than that of $|t_1|$ and v: R is not entirely determined

by $\mathcal{V}_1$ or even $\mathcal{V}_1 \cup \mathcal{V}_2$. The HDE$[R, \mathcal{V}_1]$ yields $f_\infty[R, \mathcal{V}_1] = 0.849$, so that $\mathcal{D}[R, \mathcal{V}_1]$ is incorrect. The HDE$[R, \mathcal{V}_1 \cup \mathcal{V}_2]$ yields $f_\infty[R, \mathcal{V}_1 \cup \mathcal{V}_2] = 0.860$, so that $\mathcal{D}[R, \mathcal{V}_1 \cup \mathcal{V}_2]$ is still incorrect.

Even though $\mathcal{D}[R, \mathcal{V}_1]$ is incorrect, the MOD2 of R on $\mathcal{V}_1$ [Eq. (5)] reveals the main-order mechanism of the dependence of R. Namely, R has a very rough MOD2 on $|Z_X|$ and n_{AB}, which is consistent with the below discussion. The score analysis shows that Z_X and n_{AB} correspond respectively to $x_{i_1}^{opt}$ and $x_{i_2}^{opt}$ unambiguously (see Appendix E). In Appendix F, the robustness of Eq. (5) is confirmed.

Qualitatively, R increases when (i) $|Z_X|$ decreases or (ii) n_{AB} increases in Eq. (5) (see also Fig. 1). The microscopic mechanism of (i) and (ii) is detailed below.

(i) Decreasing $|Z_X|$ reduces the negative charge of the apical anion. This reduces the negative MP created by the apical anion and felt by the electrons in the nearby CuO_2 plane. [See Fig. 9(b) for an illustration.] This reduces the energy of the electrons in the CuO_2 plane, and also reduces the Fermi energy. (Indeed, the electrons in the CuO_2 plane are near the Fermi level, so that the Fermi level is determined by the energy of the electrons in the CuO_2 plane.) On the other hand, the empty states are less affected, and their energy does not change substantially. However, because the Fermi level is reduced as discussed above, the empty states become higher in energy relative to the Fermi level (so that ϵ_e increases). This reduces the screening from empty states, and thus, increases R.

(ii) The decrease in R with decreasing n_{AB} (increasing δ) is discussed in Ref. [19], and the microscopic mechanism is reminded here. When δ increases, the hole doping of O sites increases. This reduces the negative charge of in-plane O anions. This reduces the negative MP created by the in-plane O and felt by the nearby Cu. [See Fig. 9(c) for an illustration.] This reduces the energy of Cu3d electrons, and also reduces the Fermi energy. (Indeed, the AB band at the Fermi level has Cu3$d_{x^2-y^2}$ character.) On the other hand, the O2p electrons are less affected. However, because the Fermi level is reduced, the occupied O2p states become closer to the Fermi level (so that $|\epsilon_o|$ decreases). This increases the screening from occupied states, and thus, reduces R.

For completeness, the hole doping dependence of R is discussed in detail in Appendix G, which is summarized here. Even though R increases when $|Z_X|$ decreases as discussed above, the decrease in $|Z_X|$ also accelerates the decrease in R with decreasing n_{AB}, which is not captured by Eq. (5). This is why the three color points with the lowest R in Fig. 11(b) deviate from the linear interpolation, which is a major cause of the relatively low value of $f_{(2)}[R, \mathcal{V}_1] = 0.642$. The MOD2 in Eq. (5) may be combined with the results in Appendix G to obtain a more accurate picture of the dependence of R on $|Z_X|$ and n_{AB}.

(7) Dependence of R on $\mathcal{V}_3$. R may be entirely determined by the subset $\{\epsilon_{p_x}^X, \epsilon_o, \epsilon_e\}$ of $\mathcal{V}_3$. The HDE$[R, \mathcal{V}_3]$ yields $f_\infty[R, \mathcal{V}_3] = 0.976$, so that $\mathcal{D}[R, \mathcal{V}_3]$ is reasonably correct; also, the dependence of R on $x_{(15)}^{opt}$ is almost linear [see Fig. 11(e)]. The rHDE$[R, \mathcal{V}_3]$ yields $f_\infty[R, \{|\epsilon_{p_x}^X|, |\epsilon_o|, \epsilon_e\}] = 0.974$, so that $\mathcal{D}[R, \{|\epsilon_{p_x}^X|, |\epsilon_o|, \epsilon_e\}]$ is reasonably correct as well. Note that even though R has complex dependencies

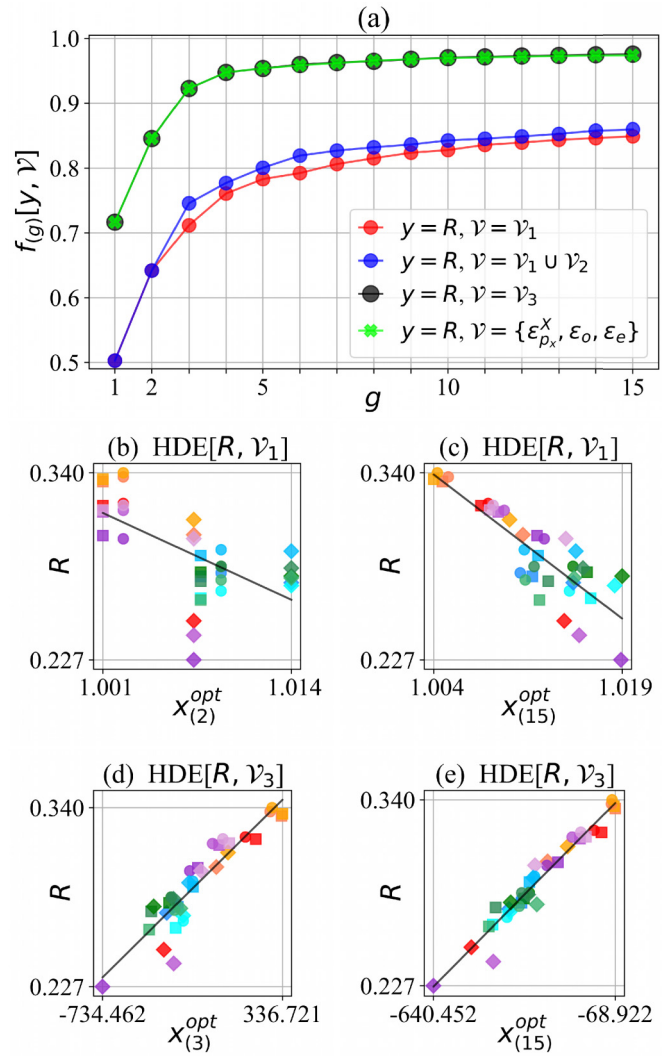

FIG. 11. Dependence of R on $\mathcal{V}_1$ and $\mathcal{V}_3$. (a) Values of $f_{(g)}[R, \mathcal{V}]$ in the HDE$[R, \mathcal{V}_1]$, the HDE$[R, \mathcal{V}_1 \cup \mathcal{V}_2]$, the HDE$[R, \mathcal{V}_3]$ and the rHDE$[R, \mathcal{V}_3]$. (b), (c) Dependence of R on $x_{(2)}^{opt}$ [corresponding to Eq. (5)] and $x_{(15)}^{opt}$ in the HDE$[R, \mathcal{V}_1]$. (d), (e) Dependence of R on $x_{(3)}^{opt}$ [corresponding to Eq. (26)] and $x_{(15)}^{opt}$ in the HDE$[R, \mathcal{V}_3]$. In panels (b)–(e), the CF that corresponds to each color point is shown in Fig. 3, and the solid black line shows the linear interpolation.

on the whole band structure *via* the polarization formula in Appendix D, Eq. (D1), the above result shows that R can be described reasonably by only three characteristic energies in the band structure.

In the following, we discuss the MOD3 of R on $\mathcal{V}_3$ instead of the MOD2 as usually done before. This is justified as follows: In the case of $|t_1|$ and v, we have $f_{(2)}[|t_1|, \mathcal{V}_3] = 0.994$ and $f_{(2)}[v, \mathcal{V}_3] = 0.984$, so that the MOD2 is accurate. However, for R, we have $f_{(2)}[y, \mathcal{V}_3] = 0.846$ but $f_{(3)}[y, \mathcal{V}_3] = 0.923$, so that the MOD3 is more accurate than the MOD2. To obtain a compromise between accuracy and simplicity, we choose to discuss the MOD3. Although the MOD3 does not describe R perfectly, the dependence of R on $x_{(3)}^{opt}$ is almost

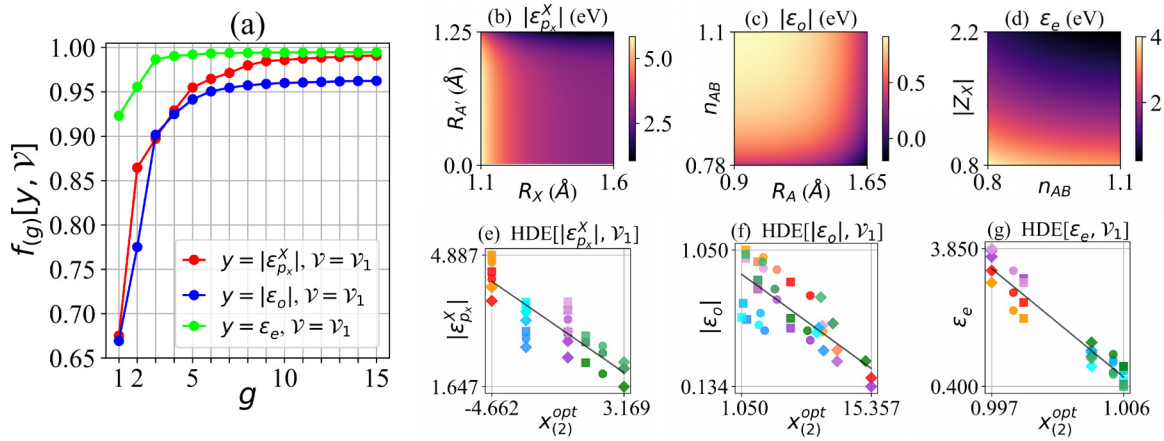

FIG. 12. Dependence of $y = |\epsilon^X_{p_x}|$, $|\epsilon_o|$, and ϵ_e on $\mathcal{V}_1$. (a) Values of $f_{(g)}[y, \mathcal{V}]$ in the HDE[$y, \mathcal{V}_1$]. (b)–(d) Representation of the MOD2 of y on $\mathcal{V}_1$ in Eqs. (27), (28), (29). (e)–(g) Dependence of y on $x^{\rm opt}_{(2)}$ which corresponds to the panels (b)–(d), respectively. The CF that corresponds to each color point is shown in Fig. 3, and the solid black line shows the linear interpolation.

linear as seen in Fig. 11(d). The MOD3 of R on $\mathcal{V}_3$ is

$$R_{\rm MOD3} = 0.309 + 0.000105 \times \left[|\epsilon^X_{p_x}|^{3.94} - 93.1|\epsilon_o|^{-1.01} - 184\epsilon_e^{-0.93}\right]. \quad (26)$$

The score analysis confirms that $|\epsilon^X_{p_x}|$, $|\epsilon_o|$, and ϵ_e correspond respectively to $x^{\rm opt}_{i_1}$, $x^{\rm opt}_{i_2}$, and $x^{\rm opt}_{i_3}$ (see Appendix E).

Qualitatively, R increases (i.e., the screening decreases) when $|\epsilon^X_{p_x}|$ increases (i.e., the occupied apical X2p_x orbital becomes farther from the Fermi level), $|\epsilon_o|$ increases (the highest occupied energy band becomes farther from the Fermi level), or ϵ_e increases (the lowest empty energy band becomes farther from the Fermi level). These three dependencies are intuitive, because the screening from a given band decreases when the band energy is farther from the Fermi level. [See the polarization formula in Appendix D, Eq. (D1).]

The three variables $\{|\epsilon^X_{p_x}|, |\epsilon_o|, \epsilon_e\}$ are entirely determined by $\mathcal{V}_1$ but not by $\mathcal{V}_2$. For y in $\{|\epsilon^X_{p_x}|, |\epsilon_o|, \epsilon_e\}$, the HDE[$y, \mathcal{V}_2$] yields $f_\infty[|\epsilon^X_{p_x}|, \mathcal{V}_2] = 0.892$, $f_\infty[|\epsilon_o|, \mathcal{V}_2] = 0.728$, and $f_\infty[\epsilon_e, \mathcal{V}_2] = 0.948$, so that $\mathcal{V}_2$ does not describe accurately $|\epsilon^X_{p_x}|$, $|\epsilon_o|$, and ϵ_e. However, the HDE[$y, \mathcal{V}_1$] yields $f_\infty[|\epsilon^X_{p_x}|, \mathcal{V}_1] = 0.991$, $f_\infty[|\epsilon_o|, \mathcal{V}_1] = 0.962$, and $f_\infty[\epsilon_e, \mathcal{V}_1] = 0.995$ [see Fig. 12(a)]. Here, we choose to express $|\epsilon^X_{p_x}|$, $|\epsilon_o|$, and ϵ_e directly as a function of $\mathcal{V}_1$ instead of $\mathcal{V}_2$ in order to obtain a more accurate expression. Thus, in the following, we discuss the dependence of $|\epsilon^X_{p_x}|$, $|\epsilon_o|$, and ϵ_e on $\mathcal{V}_1$. The MOD2s of $|\epsilon^X_{p_x}|$, $|\epsilon_o|$, and ϵ_e on $\mathcal{V}_1$ are

$$\left|\epsilon^X_{p_x}\right|_{\rm MOD2} = 2.902 - 0.287\left[R^{8.49}_{A'} - 26.5R^{-9.99}_X\right], \quad (27)$$

$$|\epsilon_o|_{\rm MOD2} = 0.934 - 0.0443\left[n^{-9.99}_{AB} + 0.151R^{8.99}_A\right], \quad (28)$$

$$\epsilon_{e\rm MOD2} = 298.062 - 295.636\left[|Z_X|^{0.01} - 0.000899n^{-5.58}_{AB}\right], \quad (29)$$

and are represented in Fig. 12. [These correspond to items (8), (9), and (10) in Fig. 5, respectively.] The score analysis confirms that $x^{\rm opt}_{i_1}$ and $x^{\rm opt}_{i_2}$ correspond to $R_{A'}$ and R_X in the dependence of $|\epsilon^X_{p_x}|$, n_{AB}, and R_A in the dependence of $|\epsilon_o|$, and Z_X and n_{AB} in the dependence of ϵ_e (see Appendix E).

On (III) and (7), let us discuss the common dependencies of R, $|\epsilon^X_{p_x}|$, $|\epsilon_o|$, and ϵ_e on $|Z_X|$ and n_{AB}. First, as for $|Z_X|$, the MODs of R and $|\epsilon^X_{p_x}|$ on $|Z_X|$ are consistent. Indeed, both $R_{\rm MOD2}$ [Eq. (5)] and $|\epsilon^X_{p_x}|_{\rm MOD2}$ [Eq. (27)] increase when $|Z_X|$ decreases; consistently, $R_{\rm MOD3}$ [Eq. (26)] increases when $|\epsilon^X_{p_x}|$ increases.

Second, as for n_{AB}, $R_{\rm MOD2}$ decreases with decreasing n_{AB} in Eq. (5). This is the result of a competition between the MODs of $|\epsilon_o|$ and ϵ_e on n_{AB} [Eqs. (28) and (29)]. Indeed, when n_{AB} decreases, the two mechanisms (i) and (ii) occur. On the one hand, (i) $|\epsilon_o|_{\rm MOD2}$ [Eq. (28)] decreases, which contributes to decrease $R_{\rm MOD3}$ according to Eq. (26). On the other hand, (ii) $\epsilon_{e\rm MOD2}$ [Eq. (29)] increases, which contributes to increase $R_{\rm MOD3}$ according to Eq. (26). The fact that $R_{\rm MOD2}$ decreases with decreasing n_{AB} in Eq. (5) suggests that (ii) dominates over (i).

(8) Dependence of $|\epsilon^X_{p_x}|$ on $\mathcal{V}_1$. The MOD of $|\epsilon^X_{p_x}|$ on $R_{A'}$ and R_X in Eq. (27) is understood as follows: First, $|\epsilon^X_{p_x}|$ decreases when $R_{A'}$ increases. Zero $R_{A'}$ corresponds to A$'$ = ∅, whereas nonzero $R_{A'}$ corresponds to A$'$ = Hg$_{1-\delta}$Au$_\delta$. The symmetry of the primitive cell changes from A$'$ = ∅ to A$'$ = Hg$_{1-\delta}$Au$_\delta$ (see Fig. 2), and the crystalline environment changes as well. And, if we represent $|\epsilon^X_{p_x}|$ as a function of $|\epsilon^X_{p_z}|$ (in Appendix D 3), we see that $|\epsilon^X_{p_x}| \simeq k_0 + k_1|\epsilon^X_{p_z}|$ has an affine dependence on $|\epsilon^X_{p_z}|$, but the coefficients k_0 are different for A$'$ = ∅ and A$'$ = Hg$_{1-\delta}$Au$_\delta$ whereas the coefficients k_1 are nearly identical. Namely, the affine regression yields

$$\left|\epsilon^X_{p_x}\right| = 0.181 + 0.940\left|\epsilon^X_{p_z}\right| \ (A' = \varnothing), \quad (30)$$

$$\left|\epsilon^X_{p_x}\right| = -1.436 + 0.965\left|\epsilon^X_{p_z}\right| \ (A' = {\rm Hg}_{1-\delta}{\rm Au}_\delta). \quad (31)$$

Thus, from A$'$ = ∅ to A$'$ = Hg$_{1-\delta}$Au$_\delta$, $|\epsilon^X_{p_x}|$ is reduced by 1.62 eV for a given value of $|\epsilon^X_{p_z}|$. On the other hand, for A$'$ = ∅, the value of k_0 is universal irrespective of A and X. This suggests that the decrease in $|\epsilon^X_{p_x}|$ from A$'$ = ∅ to A$'$ = Hg$_{1-\delta}$Au$_\delta$ does not depend on A or X, but rather on the

presence of the A′ atom and the subsequent change in crystal structure and crystal electric field. A possible explanation is the following: For $A' = Hg_{1-\delta}Au_{\delta}$, there is a A′ atom close to the apical X (see Fig. 2), and the apical X2p orbital overlaps with the A′5d orbitals. This may cause the apical X2p_x orbital to catch antibonding X2p/A′5d character,[3] which may destabilize the apical X2p_x orbital (i.e., increase its onsite energy and thus reduce $|\epsilon^{X}_{p_x}|$). This is consistent with the isosurface of the apical X2p_x orbital in Fig. 2 for $A' = Hg_{1-\delta}Au_{\delta}$: We see that the apical X2p_x orbital has a slight A′yz/zx character near the A′ atom.

Second, $|\epsilon^{X}_{p_x}|$ increases when R_X decreases. This is consistent with Ref. [20]: When a decreases, the occupied bands in the M space (including apical X bands) become farther from the Fermi level [see Ref. [20], Figs. 3(h)–3(k)]. This is because the negative MP created by the in-plane O and felt by the Cu increases [as illustrated in Fig. 9(c)], so that the energy of the Cu3$d_{x^2-y^2}$ increases (and this shifts the Fermi level upward), whereas the apical X orbitals are less affected. And, the decrease in R_X contributes to decrease a, as discussed previously.

(9) Dependence of $|\epsilon_o|$ on $\mathcal{V}_1$. The MOD of $|\epsilon_o|$ on n_{AB} and R_A in Eq. (28) is understood as follows: First, $|\epsilon_o|$ decreases when n_{AB} decreases (i.e., the hole doping increases). This is consistent with Ref. [19], and the cause is interpreted as a rigid shift of the Fermi level upon hole doping. Because the partially filled AB band is the only band at the Fermi level, increasing the hole doping reduces the number of electrons in the AB band, which shifts the Fermi level downward. As a result, the energy of occupied bands relative to the Fermi level increases. These include the highest occupied band outside the AB band, whose energy is ϵ_o. Thus, $\epsilon_o < 0$ increases, so $|\epsilon_o|$ decreases.

Second, $|\epsilon_o|$ increases when R_A decreases. The microscopic mechanism is discussed in detail in Appendix H.

(10) Dependence of ϵ_e on $\mathcal{V}_1$. The MOD of ϵ_e on $|Z_X|$ and n_{AB} in Eq. (29) is understood as follows: First, ϵ_e increases when $|Z_X|$ decreases. The mechanism was summarized in Sec. II [see (IV), item (i)], and is detailed here. Apical X anions with the negative charge Z_X emit a negative MP that increases the energy of surrounding electrons, in particular those in the M bands, because the Cu and in-plane O atoms are in the vicinity of apical X. Reducing $|Z_X|$ reduces the negative MP from apical X felt by the Cu and in-plane O. [See Fig. 9(b) for an illustration.] This reduces the energy of M bands. The Fermi energy is also reduced, because it is determined by the partially filled AB band which is in the M space. Thus, the empty bands become farther from the Fermi level. This is consistent with results on Hg1223 in Ref. [20] (see, e.g., Appendix E1).

Second, ϵ_e increases when n_{AB} decreases. This is because (i) the Fermi level is shifted downward when n_{AB} decreases because of the hole doping of the AB band [as discussed in (9)]. As a consequence, the empty bands become farther from the Fermi level, and thus ϵ_e increases. In addition, (ii) when n_{AB} decreases, the M bands are stabilized, which further shifts the M bands and thus the Fermi level downward. This is because the hole doping of in-plane O increases when n_{AB} decreases as mentioned before: This reduces the negative charge of the in-plane O ions, and thus the negative MP created by the in-plane O ions. [See Fig. 9(c) for an illustration.] This stabilizes the M bands, which increases the energy gap between M bands and empty bands.

[3]Note that the X2p_x orbital that is considered here is a maximally localized Wannier orbital, whose character may be slightly different from the purely atomic p_x character.

V. DISCUSSION

Here, we discuss the universality and accuracy of the obtained expressions of $|t_1|$, v, and R. Also, we propose prescriptions to optimize T_c^{opt} in future design of superconducting cuprates and akin materials.

A. Universality and accuracy of the expressions of AB Hamiltonian parameters

The CF dependencies of $|t_1|$, v, and R obtained in this paper offer reliable guidelines for design of SC cuprates and akin compounds, if we assume that (A) the training set considered in this paper is representative of the diversity in single-layer cuprates, and (B) the CF dependence of the AB Hamiltonian parameters $|t_1|$ and u is correctly captured (at least qualitatively) by the GGA + cRPA version of MACE employed in this paper.

(A) is supported in Appendix B; below, we support (B). The detailed discussion on (B) is necessary, because the GGA + cRPA is the simplest level of the MACE scheme; more sophisticated versions of MACE have been employed in previous works, such as the constrained *GW* (c*GW*) supplemented with self-interaction correction (SIC) at the c*GW*-SIC level [17] and level renormalization feedback (LRFB) at the c*GW*-SIC + LRFB level [18]. Equation (1) was determined by solving AB Hamiltonians at the c*GW*-SIC + LRFB level [19].

The GGA + cRPA is expected to capture correctly the qualitative materials dependence of $|t_1|$ and u [19,20] while avoiding the complexity of the c*GW*-SIC + LRFB calculation. For instance, u is smaller in Bi2201 compared with Bi2212 at the c*GW*-SIC + LRFB level, and this result is reproduced qualitatively by the GGA + cRPA [19]. In addition, in Hg1223, the qualitative pressure dependence of $|t_1|$ and u is captured by the GGA + cRPA [20].

Still, it should be noted that the accurate prediction of the materials dependence of T_c^{opt} cannot be done by considering the materials dependent $|t_1|$ and u at the GGA + cRPA level (the values in Fig. 4). Indeed, the GGA + cRPA has a limitation at the quantitative level. For instance, in Bi2201 and Bi2212, u is underestimated at the GGA + cRPA level, and the difference between the values of u ($|t_1|$) at the GGA + cRPA and c*GW*-SIC + LRFB levels is around 10% (5%) [19]. The uncertainty on u may cause a significant uncertainty on F_{SC} and thus T_c^{opt}, because T_c^{opt} strongly depends on u *via* F_{SC} in Eq. (1), especially when u is located in the weak-coupling region $u \simeq 6.5$–8.0 [14]. Thus, the quantitative improvement from the GGA + cRPA level to the c*GW*-SIC + LRFB level is required to tackle the accurate prediction of $T_c^{opt} \simeq 0.16|t_1|F_{SC}$ [14] from the values of $|t_1|$ and u.

Nonetheless, we restrict to the simplest GGA + cRPA level in the present paper, because (i) the cGW-SIC + LRFB calculation is complex and computationally expensive, and (ii) the GGA + cRPA is expected to capture correctly the materials dependence of the AB Hamiltonian at least qualitatively [19,20], as discussed above. Quantitative prediction of the CF dependence of $T_c^{\rm opt}$ by using Eq. (1) requires the CF dependence of the cGW-SIC + LRFB result, which is left for future studies.

Also, note that some of the results obtained in this paper are independent of the restriction to the GGA + cRPA level and are expected to remain valid for the AB Hamiltonian at the cGW-SIC + LRFB level. This is the case of the items (3), (5), (6), and (7) in Fig. 5. For instance, the dependencies of R and v on $\mathcal{V}_3$ in Eqs. (26) and (24) make complete abstraction of the level of the electronic structure from which the AB Hamiltonian is calculated (GGA in the case of GGA + cRPA, or GW + LRFB [18] in the case of cGW-SIC + LRFB). Thus, the dependencies of R and v on $\mathcal{V}_3$ in Eqs. (26) and (24) are expected to be rather universal. (The detailed dependencies of $|t_{xp}|$, ΔE_{xp}, $|\epsilon^{\rm X}_{p_x}|$, $|\epsilon_o|$, and ϵ_e at the GW + LRFB level on the CF are left for future studies.) However, on $|t_1|$, the dependence on $\mathcal{V}_3$ in Eq. (17) does not take into account the removal of exchange-correlation double counting [16] that is done at the cGW-SIC + LRFB level. This corrects $|t_1|$ by a term which is materials dependent [19,20]. (The detailed materials dependence of this term is left for future studies.)

B. Prescriptions to optimize $T_c^{\rm opt}$

The experimentally known values of $T_c^{\rm opt}$ at ambient pressure reach up to 94 K [4] in single-layer cuprates, and the highest known value is 138 K [2] in the triple-layer cuprate Hg1223 [2,3]. Under pressure, $T_c^{\rm opt}$ increases from 94 to 118 K in the single-layer cuprate $HgBa_2CuO_4$ and from 138 K to 164–166 K in the trilayer cuprate Hg1223 [2,3].

To optimize $T_c^{\rm opt}$ beyond the above values in future design of superconducting cuprates, we propose prescriptions by considering the MODs given in Sec. II. Because our calculations are restricted to single-layer cuprates at ambient pressure, quantitative predictions from the MODs are restricted to these compounds as well. However, the qualitative CF dependence of $|t_1|$ and u may remain valid in single-layer cuprates under pressure and trilayer cuprates, as discussed below.

The overall strategy is to maximize $|t_1|$ while keeping $u = vR/|t_1|$ as close as possible to its optimal value $u_{\rm opt} \simeq 8.0$–8.5. Indeed, $T_c^{\rm opt}$ increases with both $|t_1|$ and $F_{\rm SC}$ in Eq. (1), and the universal u dependence of $F_{\rm SC}$ has a maximum at $u_{\rm opt}$ (see Appendix A). Thus, we should satisfy the criterion

$$vR/|t_1| \simeq 8.5. \tag{32}$$

To do so, the values of $|t_1|$, v, and R may be tuned rather independently if we consider their distinct dependencies on the CF. For instance, $Z_{\rm A}$ appears in the MOD of v [Eq. (4)] but not in that of $|t_1|$ [Eq. (2)], so that tuning $Z_{\rm A}$ allows us to tune v without affecting $|t_1|$ substantially. Note that, in the particular case of X = F, Cl, the decrease in R with decreasing $n_{\rm AB}$ is sharper compared with X = O (see Appendix G), which is an effect beyond the MOD2 in Eq. (5). Possible implications on superconducting properties are discussed in Appendix G.

For single-layer cuprates at ambient pressure, our result predicts an upper bound $T_{c,\max}^{\rm opt} \simeq 140$ K for $T_c^{\rm opt}$, which is above the maximal known value $T_c^{\rm opt} = 94$ K [4]. Indeed, according to Eq. (2), the maximal value of $|t_1|$ is $|t_1|_{\max} \simeq 0.53$ eV (0.58 eV if we consider $g = 15$ instead of $g = 2$), and the maximal value of $F_{\rm SC}$ is $F_{\rm SC,max} \simeq 0.13$ at $u = u_{\rm opt}$ in the universal u dependence of $F_{\rm SC}$ [14]. Thus, for $|t_1|_{\max} \simeq 0.58$ eV, the maximal value of $T_c^{\rm opt}$ is $T_{c,\max}^{\rm opt} = 0.16|t_1|_{\max} F_{\rm SC,max} \simeq 0.012$ eV (140 K) according to Eq. (1).

The upper bound $T_{c,\max}^{\rm opt} \simeq 140$ K at ambient pressure is optimistic, because $|t_1|_{\max}$ corresponds to $R_{\rm X} = R_{\rm A} = 0$, which cannot be reached in experiment. Nonetheless, reducing $R_{\rm X}$ or $R_{\rm A}$ may allow us to make $|t_1|$ as close as possible to $|t_1|_{\max}$. This can be done by replacing A or X by an isovalent atom with a smaller ionic radius: This contributes to reduce the distances between atoms in the crystal and thus a according to Eq. (20) (by mimicking the effects of physical P_a at the chemical level).

Also, the upper bound $T_{c,\max}^{\rm opt}$ may increase if we apply physical pressure. Indeed, $|t_1|$ may exceed the upper bound $|t_1|_{\max}$ at ambient pressure. For instance, in Hg1223, $|t_1| \simeq 0.57$–0.60 eV at $P = 30$ GPa, and $|t_1| \simeq 0.62$–0.67 eV at $P = 60$ GPa [20], which is above $|t_1|_{\max} \simeq 0.58$ eV. Note that the increase in $|t_1|$ eventually causes a rapid decrease in u and thus $F_{\rm SC}$ when u falls into the weak-coupling region [20]; this may be countered by tuning $Z_{\rm A}$ and $Z_{\rm X}$ to increase v and R, so that Eq. (32) is still satisfied at high pressure.

Finally, the physical dependencies of $|t_1|$ and u on the ionic radii and charges may remain valid at least qualitatively in trilayer cuprates, in which the highest $T_c^{\rm opt}$ has been observed. On $|t_1|$, the mechanism is the following: Reducing the ionic radii reduces chemical pressure and thus interatomic distances, increasing the overlap between neighboring orbitals. On the dependence of v and R and thus u on ionic charges, the mechanism proposed in this paper involves the Madelung potential created by ions that surround Cu and O. In trilayer cuprates and even other materials, the mechanisms are expected to be similar, even though Eqs. (2) and (3) may not be valid quantitatively.

VI. CONCLUSION

We have proposed the universal CF dependence of the AB Hamiltonian parameters $|t_1|$ and u in single-layer cuprates, by proposing the HDE procedure and applying it to analyze the results of *ab initio* calculations of the AB Hamiltonian for various single-layer cuprates. The results and especially the MODs given in Sec. II provide insights to optimize $T_c^{\rm opt}$ in future design of superconducting cuprates as proposed in Sec. V.

The qualitative insights obtained in this paper may also be useful for design of other superconducting materials whose crystal structure is similar to that of cuprates, namely, other strongly correlated electron materials whose low-energy physics may be described by the AB Hamiltonian on the two-dimensional square lattice. These include nickelates [38] and the recently proposed Ag- and Pd-based compounds [24,39].

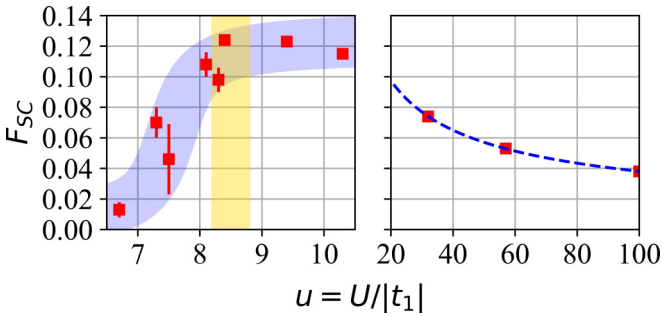

FIG. 13. Dependence of the superconducting order parameter F_{SC} on $u = U/|t_1|$, from Ref. [14]. The values shown by the red squares and error bars are taken from Figs. 10 and 21 in Ref. [14] (at hole doping $\delta = 0.167$). The blue area shows the rough shape of the u dependence of F_{SC} for $u \lesssim 10.5$. The golden vertical bar shows the optimal regime ($u \simeq u_{opt} \simeq 8.5$) in which F_{SC} reaches its maximum. The dashed blue curve shows the scaling $F_{SC} \simeq 0.56u^{-0.58}$ in the limit of large u.

More generally, the gMACE + HDE procedure employed in this paper may be used for design of strongly correlated electron materials including those which do not have superconducting properties. Even more generally, the HDE procedure offers a general platform to extract dependencies between any physical quantities y and x_i, regardless their physical meaning. These quantities may be either calculated within a theoretical framework or measured in an experiment.

ACKNOWLEDGMENTS

This work was done under the Special Postdoctoral Researcher Program at RIKEN. Part of the figures were drawn by using the software VESTA [40]. We thank Masatoshi Imada, Youhei Yamaji, and Shiro Sakai for discussions during the preliminary phase of the project.

APPENDIX A: DEPENDENCE OF SUPERCONDUCTING ORDER PARAMETER ON $u = U/|t_1|$

In Sec. I, we mention the u dependence of F_{SC} from Ref. [14]. Here, as a complement, we show this dependence in Fig. 13. In addition, we discuss the focus on $|t_1|$ and u in the scope of this paper.

The correlation between AB Hamiltonian parameters and T_c have been studied in previous experimental and theoretical works [14,41–50]. However, if we solve the full AB Hamiltonian with the nth-nearest-neighbor hopping t_n and off-site Coulomb interaction V_n up to $n = 9$ and deduce F_{SC}, then F_{SC} has little dependence on effective parameters other than u as discussed in detail in Ref. [14]. That is why T_c^{opt} in Eq. (1) is mainly determined by $|t_1|$ and u, and we focus on these two parameters in this paper.

For completeness, the dependence of F_{SC} on other effective parameters has also been examined in Ref. [14]. For instance, the dependence of F_{SC} on $|t_2/t_1|$ is small at least for $|t_2/t_1| \leqslant 0.25$ (the variation in F_{SC} does not exceed $\simeq$10%, as seen in Fig. 17 in Ref. [14]). The range of values $|t_2/t_1| \leqslant 0.25$ includes the range of *ab initio* values of $|t_2/t_1| = 0.145$–0.220 obtained in this paper. The nonzero *ab initio* off-site interaction V_n reduces the value of F_{SC} with respect to the case in which only U is considered (see, e.g., Ref. [51], Fig. 4); however, this reduction in F_{SC} has no significant materials dependence (the variation in F_{SC} in the realistic range of V_n does not exceed $\simeq$10% either, as seen in Fig. 9 in Ref. [14]).

TABLE I. List of experimentally confirmed superconducting materials whose chemical formula is included in the training set. T_c^{expt} is the experimental value of T_c from the reference given in the third column.

Chemical formula	T_c^{expt} (K)	Reference
$HgBa_2CuO_4$	$\simeq 94$	[4]
$HgSr_2CuO_4$	$\simeq 78$ (with Mo substitution)	[5]
La_2CuO_4	$\simeq 40$	[6]
$Sr_2CuO_2F_2$	$\simeq 46$	[8]
$Ca_{2-x}K_xCuO_2Cl_2$	$\simeq 24$	[12]

APPENDIX B: CHOICE OF THE TRAINING SET OF CUPRATES FOR THE gMACE + HDE PROCEDURE

In Sec. V, we assume that (A) the training set considered in this paper and represented in Fig. 3 is representative of the diversity in single-layer cuprates. Here, we support (A).

First, the training set includes CFs that correspond to experimentally confirmed superconducting cuprates with a diverse distribution of $T_c^{opt} \simeq$ 24–94 K, including $HgBa_2CuO_4$ which has the highest known value of $T_c^{opt} \simeq 94$ K among single-layer cuprates at ambient pressure. These cuprates are listed in Table I together with the experimental values of T_c.

Second, the *ab initio* values of $|t_1|$ and u obtained within the training set reproduce the diverse distribution of $|t_1|$ and u observed in realistic cuprates. We have $|t_1| \simeq 0.40$–0.55 eV and $u \simeq$ 7–10.5 in Fig. 4(a). The range of u corresponds to that observed not only in single-layer cuprates but also in multilayer cuprates [14,19,20]. In particular, the values of u correspond to the range $u \simeq 6.5$–10.5 in which F_{SC} is nonzero (see Ref. [14], Fig. 10), and thus T_c^{opt} is nonzero according to Eq. (1). This suggests the training set is a good platform to study the microscopic mechanism of the materials dependence of $|t_1|$ and u and thus T_c^{opt}.

Note that for the CFs in the training set, the M space has a similar structure, which facilitates the comparison between the variables in $\mathcal{V}_3$. Namely, the number of bands in the M space is $N_M = 17$ for all compounds in the training set. (See the band structures in Sec. S2 of the Supplemental Material [21].) In other single-layer compounds such as Bi2201 with $T_c^{opt} \simeq$ 10–40 K [1,52], we have $N_M = 23$ due to the presence of six additional bands from the BiO block layer. However, we do not consider cuprates with a Bi2201-like CF and crystal structure in order to simplify the comparison between compounds. This is not expected to weaken the generality of the result, because the diverse distribution in experimental $T_c^{opt} \simeq$ 24–94 K is already reproduced by the CFs in the training set, as mentioned above. Possible extensions of the training set to other cuprates including multilayer cuprates and Bi2201-like single-layer cuprates are left for future studies.

APPENDIX C: DETAILS ON THE HIERARCHICAL DEPENDENCE EXTRACTION PROCEDURE

The essence of the HDE was presented in Sec. III A. Here, we give more details on the HDE, including the motivation (Appendix C 1), generalities and the choice of fitness function (Appendix C 2), the wildcard operator (Appendix C 3), computational details (Appendix C 4), and numerical aspects and pitfalls (Appendix C 5).

1. Motivation of the hierarchical dependence extraction

The complete elucidation of the dependence of y on $\mathcal{V} = \{x_i\}$ requires several conditions: (a) probe the completeness of the dependence of y on $\mathcal{V}$; (b) extract the hierarchy in the dependencies of y on x_i; (c) capture the nonlinear dependence of y on x_i; (d) obtain an explicit expression of y as a function of x_i. In addition, it is desirable to (e) keep the procedure as simple as possible, and (f) preserve the sparsity in the approached expression of y (namely, the approached expression of y should depend on as few x_i as possible).

For (a), probing the completeness of the dependence of y on $\mathcal{V}$ allows us to clarify whether it is possible to construct a perfect descriptor for y as a function of x_i. This provides useful information prior to (b) and (c).

For (b), clarifying the hierarchy in the dependencies of y on x_i is necessary to pinpoint the MOD of y, namely, the variables x_i on which y depends the most. This allows us to construct an approximation of y as a function of these x_i by combining (b) and (d), which consists in $y_{\mathrm{MOD}g}$ which is defined and discussed in the main text. The MODg of y contains the principal microscopic mechanism of the dependence of y, and thus, provides useful clues in the context of materials design.

For (c), y has a nonlinear dependence on x_i in the general case. For instance, in previous works on cuprates [19,20], it has been shown that the AB Hamiltonian parameters have a nonlinear dependence on band structure variables and crystal structure variables.

For (d), the explicit expression of y as a function of x_i is necessary to understand whether y increases or decreases with increasing x_i. [Such information is not captured by (b) alone.] As mentioned above, the combination of (b) and (d) allows us to obtain $y_{\mathrm{MOD}g}$.

For (f), the sparsity allows us to facilitate the physical interpretation of the approached expression of y. Indeed, if an accurate approximation of y can be constructed from only a few x_i, the distinct contributions of the x_i to the dependence of y may be analyzed and discussed more easily.

Already existing procedures such as linear regression, polynomial regression, and also symbolic regression and sparse regression fulfill part of the above conditions (a)–(e), but not all of them. This is discussed below.

Linear regression fulfills (d) and (e), but not other points. Although it is the simplest approach (e) to obtain an explicit expression of y as a function of the x_i (d), the dependence of y on x_i is nonlinear in the general case as discussed before, so that (c) is not fulfilled.

Polynomial regression and symbolic regression fulfill (c) in addition to (d) and (e), but not other points. Polynomial regression has good approximation properties provided that y is a continuous function of the x_i: In that case, y can be approached uniformly by a polynomial of the x_i according to the Stone-Weierstrass theorem. Symbolic regression allows us to go beyond the polynomial regression without introducing assumptions on the expression of y as a function of the x_i. However, these techniques do not allow us to fulfill (a), (b), and (f) in the general case.

Sparse regression allows us to perform polynomial and symbolic regression by fulfilling (f). Regularization techniques allow us to penalize expressions of y that depend on many x_i, allowing to construct an expression of y as a function of a reduced number of x_i (f). However, they do not offer a clear framework to fulfill (a) or (b).

The HDE is designed to fulfill all conditions (a)–(f). First, the HDE allows us to probe the completeness of the dependence (a) by examining the value of $f_\infty[y, \mathcal{V}]$ in Eq. (12). Second, the HDE allows us to extract the hierarchy in dependencies of y on x_i (b) thanks to the recurrent expression of the candidate descriptor $x_{(g)}$ in Eq. (8). In the HDE$[y, \mathcal{V}]$, when we increment g, we successively add terms to x^{opt} in Eq. (11) that correspond to the higher-order dependencies of y on $\{x_i\}$ (the order increases with g). The competition between dependencies can also be analyzed by calculating the score of each variable (see Appendix E). Third, the wildcard operator in Eq. (9) encompasses polynomials of variables, so that the nonlinear dependence of y may be captured at an acceptable level of accuracy (c). (Still, note that its approximation properties are not perfect, as discussed below.) Fourth, the HDE allows us to obtain an explicit expression of y as a function of x_i (d). Fifth, the HDE procedure is simple and deterministic (e), and can be applied even if the size of the training set is relatively small (we should have $N_t \geqslant 3$). Sixth, the HDE procedure allows us to preserve the sparsity by introducing no more than one variable x_i in Eq. (11) when g is incremented.

Note that, in the HDE$[y, \mathcal{V}]$, the ratio

$$\Delta f_{(g)} = \frac{f_{(g+1)}[y, \mathcal{V}] - f_{(g)}[y, \mathcal{V}]}{f_{(g)}[y, \mathcal{V}]} \tag{C1}$$

usually decreases with increasing g, but the amplitude of the ratio

$$\Delta x^{\mathrm{opt}}_{(g)} = \frac{x^{\mathrm{opt}}_{(g+1)} - x^{\mathrm{opt}}_{(g)}}{x^{\mathrm{opt}}_{(g)}} = \zeta^{\mathrm{opt}}_{g+1} \frac{\left[x_{i^{\mathrm{opt}}_{g+1}}\right]^{\alpha^{\mathrm{opt}}_{g+1}}}{\left[x^{\mathrm{opt}}_{(g)}\right]^{\beta^{\mathrm{opt}}_{g+1}}} \tag{C2}$$

does not necessarily decrease with increasing g. Namely, when incrementing g, the effect on the *dependence* of y on x_i may be corrective (i.e., $f_{(g)}[y, \mathcal{V}]$ increases by a small amount so that $\Delta f_{(g)}$ is small), but the effect on the *amplitude* of x^{opt} may be significant: The amplitude of the additional term (which is encoded in $|\Delta x^{\mathrm{opt}}_{(g)}|$) is not necessarily small.

Note that the approximation properties of the HDE are not perfect due to the compromise between approximation and simplicity in the present framework. Namely, we enforce sparsity by allowing only one variable x_i in $\mathcal{V}$ to be introduced in $x^{\mathrm{opt}}_{(g)}$ when g is incremented, and the expression of $\Delta x^{\mathrm{opt}}_{(g)}$ is kept as simple as possible to facilitate its interpretation. On one hand, this choice does not allow us to represent all polynomials of x_i, which requires to introduce several variables at once in $x^{\mathrm{opt}}_{(g)}$ when g is incremented. On the other

hand, this allows us to obtain the hierarchy between the dependencies of y on x_i in a simple manner. In the scope of this paper, the consistency between results provided by the HDE and previous works suggests the HDE in its present form is reliable. Possible extensions of the HDE that involve more sophisticated expressions of $\Delta x^{\rm opt}_{(g)}$ are left for future studies.

Next, we discuss the physical justification of the dependence of y on x_i as defined in Eqs. (6) and (11). The HDE expression in Eq. (11) has been chosen for universality and mathematical simplicity rather than with a particular physical consideration. Indeed, the goal of the HDE is to propose an expression of y that is independent of any prior assumptions on the expression of y. The physical meaning of the expression of y is established *a posteriori* by interpreting Eq. (6). In the present paper, we demonstrate that the simple expressions in Eqs. (6) and (11) are sufficient to obtain a physically consistent picture for the dependence of $|t_1|$ and u, at least within the scope of the paper.

For (f), it should be noted that sparsity is distinct from dimensionality reduction. Dimensionality reduction techniques such as principal component analysis [53] have been used in e.g., recent studies on iron-based superconducting materials [54]. The principal component analysis allows us to reduce the size of the variable space $\mathcal{V} = \{x_i, i = 1 \ldots, N_\mathcal{V}\}$, by extracting principal components that are linear combinations of the x_i. (Details can be found in e.g., Ref. [55].) The mth principal component is denoted as

$$z^{(m)}[j] = \sum_{i=1}^{N_\mathcal{V}} v_i^{(m)} x_i[j]. \tag{C3}$$

The principal components are sufficient to reproduce the diverse distribution of values of x_i in the training set, so that an approached expression of y may be constructed as a function of a few $z^{(m)}$. However, in Eq. (C3), the weights $v_i^{(m)}$ may be nonzero for many variables x_i in the general case. Thus, if we try to express y as a function of $\{z^{(m)}, m\}$ instead of $\{x_i, i = 1, \ldots, N_\mathcal{V}\}$, the expression of y will depend on a few $z^{(m)}$ but may depend on many x_i, so that the sparsity is not preserved. To interpret the expression of y, we discuss the dependence of y on the x_i rather than on the $z^{(m)}$ because the variables x_i have a physical meaning (see Sec. III C).

2. General problem and comments on the choice of fitness function

The general scope of the HDE is the following. Starting from y and the variable space $\mathcal{V} = \{x_i, i = 1, \ldots, N_\mathcal{V}\}$, we define the candidate descriptor space

$$\mathcal{C}_\mathcal{V} = \{x(\mathbf{p}), \mathbf{p}\}. \tag{C4}$$

The elements $x(\mathbf{p})$ of $\mathcal{C}_\mathcal{V}$ are called "candidate descriptors," and are functions of the x_i. The analytic expression of $x(\mathbf{p})$ depends on variational parameters which are encoded in the vector $\mathbf{p}$. [In this paper, the expression of $x(\mathbf{p})$ is given in Eqs. (7) and (8) and we have $\mathbf{p}_g = (i_g, \alpha_g, \beta_g, \zeta_g)$ at generation g and $\mathbf{p}_1 = (i_1, \alpha_1)$.] The general problem consists in finding $\mathbf{p}^{\rm opt}$ such that $x^{\rm opt} = x(\mathbf{p}^{\rm opt})$ is the best candidate descriptor for y among the elements of $\mathcal{C}_\mathcal{V}$, i.e.,

$$f[y, x(\mathbf{p}^{\rm opt})] = \max_\mathbf{p} f[y, x(\mathbf{p})], \tag{C5}$$

where the fitness function f describes how well y is described by $x(\mathbf{p}^{\rm opt})$. [In this paper, we consider the fitness function in Eq. (10), and $f[y, x_{(g)}]$ is optimized at each generation g.]

The definition of the fitness function in Eq. (10) uses the Pearson correlation coefficient $\rho(y, x)$ between two variables y and x, whose definition is reminded below. The variables y and x are represented by two data samples $\{y[j],\ j = 1, \ldots, N_t\}$ and $\{x[j],\ j = 1, \ldots, N_t\}$, where N_t is the size of the training set. The sample mean of x, sample covariance of y and x and sample variance of x are defined as, respectively:

$$m(x) = \frac{1}{N_t} \sum_{j=1}^{N_t} x[j], \tag{C6}$$

$$c(y, x) = \frac{1}{N_t - 1} \sum_{j=1}^{N_t} [y[j] - m(y)][x[j] - m(x)], \tag{C7}$$

$$v(x) = c(x, x), \tag{C8}$$

and we calculate $\rho(y, x)$ as

$$\rho(y, x) = \frac{c(y, x)}{\sqrt{v(y)}\sqrt{v(x)}}. \tag{C9}$$

The value of $\rho(y, x)$ is a real number between -1 and $+1$. If $\rho(y, x) = 1$ [$\rho(y, x) = -1$], then x and y are perfectly correlated (anticorrelated), and there exist k_0 and k_1 such that the equation $y = k_0 + k_1 x$ is rigorously satisfied, with $k_1 > 0$ ($k_1 < 0$) if $\rho(y, x) = 1$ [$\rho(y, x) = -1$]. Also, $\rho(y, x)$ has the property

$$\rho[y, k_0 + k_1 x] = \mathrm{sgn}(k_1)\rho[y, x]. \tag{C10}$$

[In Eq. (C10), we have $k_1 \neq 0$, $\mathrm{sgn}(k_1) = 1$ if $k_1 > 0$ and $\mathrm{sgn}(k_1) = -1$ if $k_1 < 0$.]

The fitness function in Eq. (10) has the invariance property

$$f[y, k_0 + k_1 x(\mathbf{p})] = f[y, x(\mathbf{p})] \tag{C11}$$

(for $k_1 \neq 0$), which comes from Eq. (C10). Thus, $f[y, x(\mathbf{p})]$ encodes the affine dependence of y on $x(\mathbf{p})$, which is the relevant information and does not depend on the scale or order of magnitude of y and $x(\mathbf{p})$. The values of $f[y, x(\mathbf{p})]$ are between zero and one; if $f[y, x(\mathbf{p})] = 1$, there exist k_0, k_1 such that we have rigorously $y = k_0 + k_1 x(\mathbf{p})$, and $x(\mathbf{p})$ is deemed perfect. The closer $f[y, x(\mathbf{p})]$ is to 1, the more accurate the affine dependence is.

Note that even though y has an affine dependence on $x(\mathbf{p})$ when $f[y, x(\mathbf{p})] = 1$, $x(\mathbf{p})$ has a nonlinear dependence on x_i in the general case. Thus, nonlinearity in the dependence of y on x_i may be described by the above formalism.

The invariance property [Eq. (C11)] simplifies the search of the best candidate descriptor. From the viewpoint of Eq. (C11), $x(\mathbf{p})$ is equivalent to any other candidate descriptor in the equivalence class

$$\mathcal{E}[x(\mathbf{p})] = \{k_0 + k_1 x(\mathbf{p}),\ (k_1 \neq 0, k_0)\}, \tag{C12}$$

and we do not need to distinguish between elements of $\mathcal{E}[x(\mathbf{p})]$ during the optimization in Eq. (C5). The values of k_0 and k_1 are determined after $x^{\rm opt}$ has been determined, by performing the affine regression in Eq. (6).

Due to the invariance property [Eq. (C11)], the value of $f[y, x(\mathbf{p})]$ has a rather universal meaning (at least at a fixed

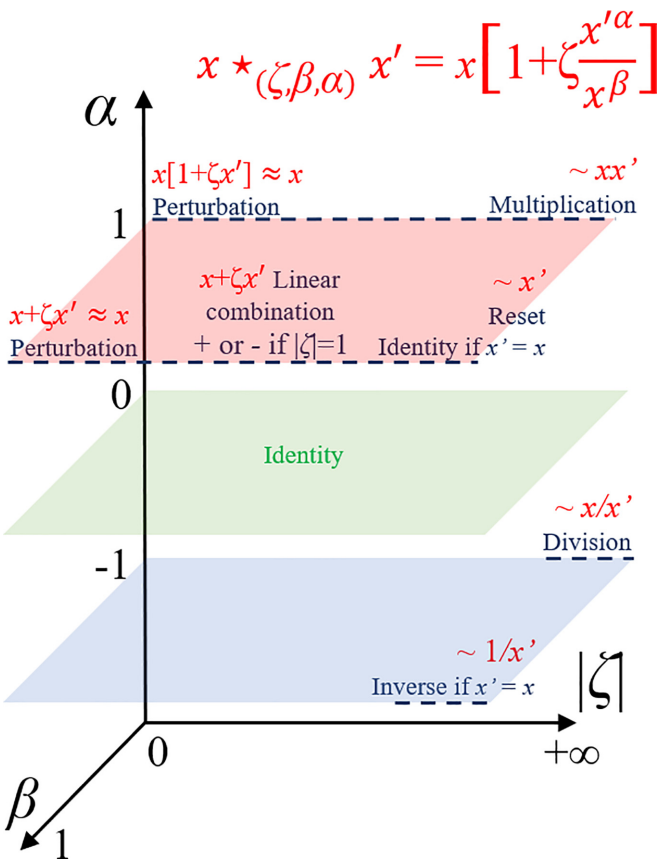

FIG. 14. Character of the wildcard operator $\star_{(\zeta,\beta,\alpha)}$ depending on the values of (ζ, β, α), when assuming the fitness function in Eq. (10). In the general case, we have $x \star_{(\zeta,\beta,\alpha)} x' = x[1 + \zeta x'^{\alpha}/x^{\beta}]$. The red, green, and blue planes correspond to $\alpha = 1$, $\alpha = 0$, and $\alpha = -1$, respectively. In Appendix C 3, we comment the character that is acquired by the wildcard operator for the values of (ζ, β, α) that correspond to the blue dashed lines. (For these values, the character is shown explicitly in the figure.) At $|\zeta| \to \infty$, we have $x \star_{(\zeta,\beta,\alpha)} x' \simeq \zeta x'^{\alpha} x^{1-\beta}$, and the factor ζ can be traced out by using the invariance property of the fitness function [Eq. (C11)].

value of N_t), which facilitates the judgment of the quality of $x(\mathbf{p})$. At $N_t = 36$, we observe empirically in Sec. IV that the description is almost perfect for $f[y, x(\mathbf{p})] \gtrsim 0.98$. [In this paper, we assume that $x(\mathbf{p})$ describes y entirely if $f_{\infty}[y, \mathcal{V}] \gtrsim 0.98$.] If $0.95 \lesssim f[y, x(\mathbf{p})] \lesssim 0.97$, the description is good, but corrective higher-order dependencies of y on x_i may be missing. If $f[y, x(\mathbf{p})] \simeq 0.6$–$0.9$, the description is rough or very rough, but $x(\mathbf{p})$ captures correctly the MOD. Typically, for the MOD2s discussed in Secs. II and IV, the values of $f_{(2)}[y, x_{(2)}^{\mathrm{opt}}]$ are above 0.6, and most of the time above 0.8–0.9. There are also particular cases in which $f_{(2)}[y, x_{(2)}^{\mathrm{opt}}] \gtrsim 0.98$, so that the MOD2 is sufficient to describe y entirely.

3. Comments on the wildcard operator

Here, we detail the properties of the wildcard operator $\star_{(\zeta,\beta,\alpha)}$ [Eq. (9)]. These properties are illustrated in Fig. 14.

(i) The wildcard operator can represent basic algebraic operations. First, multiplication and division are accounted for by

$$x \star_{(\zeta\to+\infty,\beta=0,\alpha=\pm1)} x' \propto x x'^{\pm1}. \tag{C13}$$

Indeed, for $\zeta \to +\infty$, we have

$$x \star_{(\zeta,\beta=0,\alpha)} x' = x[1 + \zeta x'^{\alpha}] \sim_{\zeta\to+\infty} \zeta x x'^{\alpha} \propto x x'^{\alpha}, \tag{C14}$$

in which the arbitrarily large yet finite factor ζ is traced out by using the invariance property of the fitness function. Furthermore, Eq. (C14) encompasses products and ratios between x and exponents of x'. Second, addition and subtraction are accounted for by

$$x \star_{(\zeta,\beta=1,\alpha=1)} x' = x + \zeta x' \tag{C15}$$

for $\zeta = \pm 1$. Furthermore, Eq. (C15) encompasses any linear combination of x and x' if $|\zeta| \neq 1$. Third, exponents and polynomials of any variable x are accounted for by

$$x \star_{(\zeta,\beta=1,\alpha)} x = x + \zeta x^{\alpha}, \tag{C16}$$

$$x \star_{(\zeta\to+\infty,\beta=1,\alpha)} x = x^{\alpha+1}. \tag{C17}$$

(ii) The wildcard operator can represent the reset and identity operators. The reset operator is defined as

$$x \star_{(\zeta\to+\infty,\beta=1,\alpha=1)} x' \propto x' \tag{C18}$$

and allows us to replace x by x'. If $x' = x$, Eq. (C18) becomes the identity operator. [The possible values of $(\zeta^{\mathrm{opt}}, \beta^{\mathrm{opt}}, \alpha^{\mathrm{opt}})$ which represent the identity operator are $(\zeta \to +\infty, \beta = 1, \alpha = 1)$ and also $(\zeta > 0, \beta, \alpha = 0)$.] The fact that the wildcard operator is able to mimic the identity operator guarantees that $x_{(g)}^{\mathrm{opt}}$ is included in the candidate descriptor space for the generation $g + 1$, and thus,

$$f\left[y, x_{(g+1)}^{\mathrm{opt}}\right] \geqslant f\left[y, x_{(g)}^{\mathrm{opt}}\right]. \tag{C19}$$

This is a desired property of the fitness function: Given the best candidate descriptor at g, the iteration from g to $g + 1$ adds the $g + 1^{\mathrm{th}}$ order dependence, and taking into account the $g + 1^{\mathrm{th}}$ order dependence necessarily improves the descriptor.

(iii) The wildcard operator can represent a perturbation of x by x'. For small $|\zeta|$ and $\alpha = 1$, we have

$$\left|\zeta \frac{x'}{x^{\beta}}\right| \ll 1 \tag{C20}$$

in Eq. (9), so that $x \star_{(\zeta,\beta,\alpha=1)} x'$ yields x corrected by a small term that depends on x'.

4. Computational details of the hierarchical dependence extraction procedure

Here, we give computational details of the procedure that is employed to obtain $x_{(N)}^{\mathrm{opt}}$ in Eq. (11). First, we use a finite number N of generations. In this paper, we use $N = 15$ in Sec. IV C, and we check that the fitness function varies slowly with increasing N at $N \simeq 15$, so that it is reasonable to stop at $N = 15$.

At fixed g, the variational parameters are $\mathbf{p}_g = (\alpha_g, \beta_g, \zeta_g, i_g)$. We determine $\mathbf{p}_g^{\mathrm{opt}} = (\alpha_g^{\mathrm{opt}}, \beta_g^{\mathrm{opt}}, \zeta_g^{\mathrm{opt}}, i_g^{\mathrm{opt}})$, which maximizes $f[y, x_{(g)}]$ as follows. We scan $\mathcal{C}_{\mathcal{V}}$ by computing $f[y, x_{(g)}]$ for each $\mathbf{p}_g$. This allows us to avoid falling into local extrema of $f[y, x(\mathbf{p})]$, but is computationally expensive as discussed below. A possibility to reduce the

computational cost would be to prepare an initial guess for **p** and then optimize **p** by using, e.g., a gradient descent algorithm. Such extensions and the detailed study of the dependence of the result on the initial guess for **p** are left for future studies.

The discretization of $\mathbf{p}_g = (\alpha_g, \beta_g, \zeta_g, i_g)$ is done as follows: The number of indices i_g is given by the total number of variables $N_{\mathcal{V}}$ in $\mathcal{V} = \{x_i\}$. As for β, we consider only $N_\beta = 2$ values, which are $\beta = 0, 1$, this choice preserves the polyvalence of the wildcard operator as illustrated in Fig. 14. As for α and ζ, we use a finite grid of values for the optimization of the fitness function. The numbers of values of α and ζ in the grid are denoted as N_α and N_ζ, respectively. The computational cost increases rapidly with N_α and N_ζ: At fixed g, the number of candidate descriptors to be evaluated is $N_{\mathcal{V}}N_\alpha$ if $g = 1$ and $N_{\mathcal{V}}N_\alpha N_\beta N_\zeta$ if $g \geqslant 2$. In practice, we reduce the computational cost by employing a multistep optimization, as detailed in (i) and (ii) below.

(i) First, we optimize (α_g, ζ_g) on a coarse grid together with (β_g, i_g). The coarse grid is the following:

$$\alpha_g^{(1)} = n, \quad n = -9, \ldots, 9, \tag{C21}$$

$$\zeta_g^{(1)} = m10^{m'}, \quad m, m' = -9, \ldots, 9 \ (m \neq 0). \tag{C22}$$

The optimized values are denoted as $(\alpha_g^{\mathrm{opt}(1)}, \beta_g^{\mathrm{opt}}, \zeta_g^{\mathrm{opt}(1)}, i_g^{\mathrm{opt}})$.

(ii) Then, we keep the values of β_g^{opt} and i_g^{opt} that were determined in (i), and we refine the optimization of (α_g, ζ_g) on a finer grid, which is constructed iteratively as follows: We introduce the iteration index j, starting from $j = 2$. The fine grid is built around $\alpha_g^{\mathrm{opt}(j-1)}, \zeta_g^{\mathrm{opt}(j-1)}$:

$$\alpha_g^{(j)} = \alpha_g^{\mathrm{opt}(j-1)}[1 + 0.1n], \quad n = -9, \ldots, 9, \tag{C23}$$

$$\zeta_g^{(j)} = \zeta_g^{\mathrm{opt}(j-1)}[1 + 0.1m], \quad m = -9, \ldots, 9. \tag{C24}$$

We optimize the fitness function to obtain $\alpha_g^{\mathrm{opt}(j)}, \zeta_g^{\mathrm{opt}(j)}$. We iterate up to $j = 3$.

Note that the grid includes values of $|\alpha|$ from $|\alpha|_{\min} = 0.01$ to $|\alpha|_{\max} = 9.99$ and values of $|\zeta|$ from $|\zeta|_{\min} = 0.01 \times 10^{-9}$ to $|\zeta|_{\max} = 9.99 \times 10^9$ if we iterate up to $j = 3$. Also, note that the score in Appendix E is calculated on the coarse grid, because i_g is optimized on the coarse grid.

5. Numerical aspects and pitfalls

Here, we discuss a few limitations and numerical pitfalls in the HDE procedure. First, note that if $\beta > 0$, the values of x [or $x_{(g-1)}^{\mathrm{opt}}$ in Eq. (8)] must be nonzero when calculating $x \star_{(\zeta,\beta,\alpha)} x' = x[1 + \zeta x'^{\alpha}/x^{\beta}]$. Namely, if the variable x is represented by the sample $\{x[j], \ j = 1, \ldots, N_t\}$ introduced in Appendix C 2, we should have $x[j] \neq 0$ for all j. This implies $x_i[j] \neq 0$ for all x_i in $\mathcal{V}$, because we have $x_{(1)}^{\mathrm{opt}} = x_{i_1}^{\alpha_1}$: If x_{i_1} has zero values, then $x_{(1)}^{\mathrm{opt}}$ will have zero values if $\alpha_1 > 0$, or $x_{(1)}^{\mathrm{opt}}$ will diverge if $\alpha_1 < 0$. In our calculations, some of the variables x_i in $\mathcal{V}$ have zero values: For instance, in $\mathcal{V}_1$, the variable $R_{\mathrm{A}'}$ has zero values if $\mathrm{A}' = \varnothing$. Thus, we introduce an infinitesimal offset $\delta_{\mathrm{off}} > 0$ in variables that have zero values. (Namely, we replace $x_i[j]$ by $x_i[j] + \delta_{\mathrm{off}}$.) We choose the value $\delta_{\mathrm{off}} = 10^{-4}$, which is small enough to be negligible with respect to the difference between values of $x_i[j]$ but large enough so that $|\zeta|_{\max}\delta_{\mathrm{off}} \gg 1$ and $|\zeta|_{\min}/\delta_{\mathrm{off}} \ll 1$. (The condition $|\zeta|_{\max}\delta_{\mathrm{off}} \gg 1$ is necessary to have $1 + \zeta x^\alpha \sim \zeta x^\alpha$ when $|\zeta| \to |\zeta|_{\max}$, so that the character of the wildcard operator at $|\zeta| \to +\infty$ shown in Fig. 14 is valid.)

Second, in some particular cases, there may be no global maximum in the α dependence of $f[y, x_{(g)}]$, e.g., $f[y, x_{i_1}^{\alpha,\mathrm{opt}}]$ at $g = 1$. Namely, the value of $|\alpha|$ becomes arbitrary large when attempting to maximize $f[y, x^\alpha]$. This is discussed in Sec. S3 of the Supplemental Material [21]. To avoid the divergence of $|\alpha|$, we consider a maximal value $|\alpha|_{\max}$ for $|\alpha|$ in practice, and if $f[y, x^\alpha]$ is maximal for α^{opt} such that $|\alpha^{\mathrm{opt}}| = |\alpha|_{\max}$, then we assume $|\alpha^{\mathrm{opt}}| = |\alpha|_{\max}$. In this paper, we have $|\alpha|_{\max} = 9.99$ by employing the grids in Eqs. (C21) and (C23) up to $j = 3$. Note that $|\alpha| = |\alpha|_{\max}$ is sometimes reached in the MODs discussed in this paper [see, e.g., the dependence of $|t_1|$ on R_{X} in Eq. (2)].

Third, the maximal value of $f[y, x_{(g)}]$ may be obtained for several candidate descriptors with the same i_g but different α_g. These are in the same equivalence class [Eq. (C12)], but this is nontrivial. This is discussed in Sec. S4 of the Supplemental Material [21]. In this case, we choose to apply the following convention: We choose α such that $|\alpha|$ is minimal and $\alpha > 0$. (Then, if $g = 2$ and if there are several values of ζ for which the fitness function has the same value, we choose ζ such that $|\zeta|$ is minimal and $\zeta > 0$.) In practice, if we use this convention, the optimization of α on the coarse grid then on the fine grid yields $\alpha = |\alpha|_{\min} = 0.01$.

A concrete example is $|Z_{\mathrm{X}}|$, whose value is either 1 and 2. In this case, $f[R, |Z_{\mathrm{X}}|^\alpha] = 0.503$ for all values of $\alpha \neq 0$. This is why we have, e.g., $\alpha_1 = 0.01$ in Eq. (5). Note that this choice should not affect the physical interpretation of the MOD of R on $|Z_{\mathrm{X}}|$ in Eq. (5). The MOD1 of R on $\mathcal{V}_1$ is

$$R = 4.22 - 3.91|Z_{\mathrm{X}}|^{0.01}, \tag{C25}$$

and if we choose another convention, e.g., we choose α such that $|\alpha|$ is minimal and $\alpha < 0$, then we obtain

$$R = -3.63 + 3.94|Z_{\mathrm{X}}|^{-0.01}, \tag{C26}$$

so that the physical interpretation of the MOD1 (R increases when $|Z_{\mathrm{X}}|$ decreases) remains the same irrespective of the selected convention. The relatively large values of $|k_0|$ and $|k_1|$ in the above MOD1 and in the MOD2 [Eq. (5)] with respect to the *ab initio* values of $R \simeq 0.22$–0.34 are a consequence of the small value of $|\alpha|$. Still, the range of *ab initio* values of R is correctly reproduced by the MOD2, as seen in Fig. 1.

APPENDIX D: DETAILS OF THE gMACE PROCEDURE

1. Computational details of the gMACE procedure

Here, we give computational details of the *ab initio* calculations in the gMACE procedure. We also detail the procedure that is employed to construct the AB maximally localized Wannier orbital.

Structural optimization and DFT calculation. We use QUANTUM ESPRESSO [56,57] and optimized norm-conserving Vanderbilt pseudopotentials [58,59] with the GGA-PBE functional [25]. To model hole doping, we use the virtual crystal approximation [31] as done in Refs. [19,20] and as mentioned in the main text: The pseudopotential of A or A′ cation is

TABLE II. Primitive vectors **a**, **b**, **c** of the Bravais lattice and atomic positions in Cartesian coordinates. We consider $c_\perp = 0$ if $A' = Hg_{1-\delta}Au_\delta$ and $c_\perp = a/2$ if $A' = \varnothing$. The atomic positions are entirely determined by a, c, $c_\perp$, d_A^z, and d_X^z. Note that there are two O atoms, two A atoms and two X atoms in the unit cell. The first and second O atoms in the unit cell are denoted as O and O′, respectively.

	x	y	z
a	a	0	0
b	0	a	0
c	$c_\perp$	$c_\perp$	c
Cu	0	0	0
O	$a/2$	0	0
O′	0	$a/2$	0
A	$a/2$	$a/2$	$\pm d_A^z$
X	0	0	$\pm d_X^z$
A′	0	0	$c/2$

interpolated with that of the chemical element whose atomic number is that of A or A′ minus one. We use a plane-wave cutoff of 100 Ry for the wave functions. The full Brillouin zone is sampled by using a k-point grid of size $8 \times 8 \times 8$ in the structural optimization, $12 \times 12 \times 12$ in the self-consistent DFT calculation, and $6 \times 6 \times 6$ in the non-self-consistent DFT calculation ($8 \times 8 \times 4$ if $A' = Hg_{1-\delta}Au_\delta$). We use a Fermi-Dirac smearing of 0.002 Ry.

Crystal parameters and symmetry. The primitive vectors of the Bravais lattice and the positions of atoms are given in Table II. During the structural optimization, we optimize the values of a, c, d_A^z, and d_X^z altogether. Then, we use the optimized values in the self-consistent and non-self-consistent DFT calculations.

For completeness, note that the high-symmetry structure that is considered in Table II may be lowered in experiment. Namely, atoms may undergo displacements around their ideal positions listed in Table II, which creates incommensurate modulations in the crystal structure. The origin of these displacements is the nonideal Goldschmidt tolerance factor of the perovskite-like structure of the cuprate (i.e., the ratio between the radii of atoms in the crystal is not ideal), which causes a geometric mismatch between the block layer and the CuO_2 plane. This happens, e.g., in the case of Bi2201 [60,61]. Also, in the case of $Sr_2CuO_2F_2$, the F atoms are distorted, and the doping may introduce excess F at different positions [8].

It is possible to account for the structural distortion in a simplified manner, by considering a distortion that is restricted to the unit cell and ignores the incommensurate character of the distortion, as done in Ref. [19] in the case of Bi2201. (See Appendix C of Ref. [19].) However, in Bi2201, the effect of the distortion $|t_1|$ and u is small [19]: If we compare the AB Hamiltonian obtained from the crystal structure with and without distortion, $|t_1|$ does not vary and U varies by no more than 3%. In the present paper, for simplicity, we do not consider the distortion and always assume the ideal atomic positions in Table II.

AB orbital. Here, we detail the procedure to construct the AB maximally localized Wannier orbital. We use the RESPACK code [19,62]. The initial guess consists in an atomic $Cu3d_{x^2-y^2}$ orbital centered on the Cu atom in the unit cell. In the outer window, the spillage functional [29] is minimized to obtain the AB subspace and band dispersion, then the spread functional [28] is minimized to obtain the AB orbital. In previous works [17–20], the outer window that is used to construct the AB orbital consists in the M space, from which the N_B lowest bands are excluded to avoid catching the $Cu3d_{x^2-y^2}/O2p_\sigma$ bonding character. However, the characteristics of the AB orbital and values of AB Hamiltonian parameters depend slightly on N_B: For instance, the AB Hamiltonian parameters may vary by a few percent if N_B varies by two or three [17,19]. Although this small dependence of the AB orbital on N_B does not change the physics of the AB Hamiltonian, it may prevent the very accurate comparison between AB Hamiltonians that are derived from different CFs.

To estimate accurately the materials dependence of the AB Hamiltonian, we employ a modified procedure to construct the AB orbital, in which we remove the bonding character without excluding the bonding bands from the outer window. This procedure is based on the antibonding-bonding transformation [24], and is described below. We use the whole M space as the outer window, and we consider three initial guesses for the maximally localized Wannier orbitals: The $Cu3d_{x^2-y^2}$ atomic orbital centered on Cu, and the $O2p_\sigma$ ($O'2p_\sigma$) orbital centered on O (O′). Then, we minimize the spillage functional [29] to obtain the band dispersion that corresponds to the three orbitals with $Cu3d_{x^2-y^2}$ and $O/O'2p_\sigma$ character. This band dispersion consists in the partly filled AB band plus the two fully occupied bonding bands; see, e.g., Fig. 8 in Ref. [19] for an illustration. Then, we discard the bonding band dispersion and we keep only the AB band, which contains the AB subspace and has been determined without introducing the dependence on N_B. Finally, we obtain the AB orbital from the AB band by considering the AB band as the outer window, reinitializing the $Cu3d_{x^2-y^2}$ initial guess and minimizing the spread functional [28]. (Note that the antibonding-bonding transformation allows us to derive a three-orbital Hamiltonian that includes the AB orbital and two bonding orbitals; here, contrary to the antibonding-bonding transformation, we discard the bonding subspace.)

Polarization. We use the RESPACK code [19,62]. The cRPA polarization at zero frequency is expressed as [62]

$$[\chi_{\mathrm{H}}]_{GG'}(q) = -\frac{4}{N_k}\sum_k \sum_{n_u}^{\text{empty}} \sum_{n_o}^{\text{occupied}} \left(1 - T_{n_o k} T_{n_u k+q}\right) \times \frac{M^G_{n_o,n_u}(k+q,k)\left[M^{G'}_{n_o,n_u}(k+q,k)\right]^*}{\Delta_{n_o,n_u}(k,q) - i\eta}, \quad \text{(D1)}$$

with

$$\Delta_{n_o,n_u}(k,q) = \epsilon_{n_u k+q} - \epsilon_{n_o k}, \quad \text{(D2)}$$

and

$$M^G_{n_o,n_u}(k+q,k) = \int_\Omega dr \psi^*_{n_u k+q}(r) e^{i(q+G)r} \psi_{n_o k}(r) \quad \text{(D3)}$$

in the above equations, η is an infinitesimal positive number (we take $\eta = 0.0272$ eV), q is a wave vector in the Brillouin zone, G, G' are reciprocal-lattice vectors, nk is the Kohn-Sham one-particle state with energy ϵ_{nk} and wave function ψ_{nk}, $T_{nk} = 1$ if nk belongs to the AB band and $T_{nk} = 0$ else. The

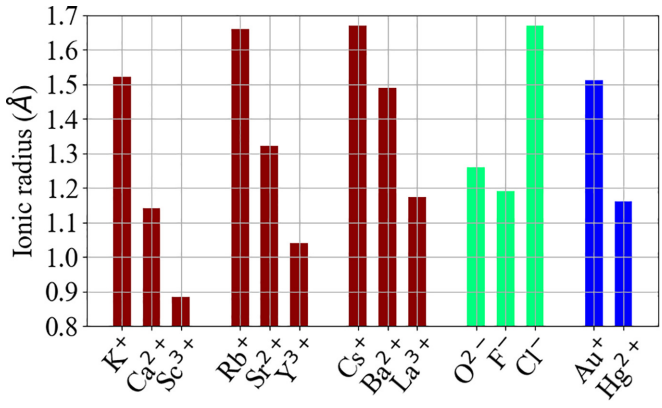

FIG. 15. Values of ionic radii and charges that are considered in this paper. The values of ionic charges are $Z_X = -2$ or -1 and $Z_A = +1$, $+2$ or $+3$. The values of R_A, R_X, and $R_{A'}$ are represented in red, green, and blue, respectively. In the case of the hole-doped compounds, the values of ionic radii and charges are interpolated according to Eqs. (D5) and (D6).

cRPA effective interaction is deduced as

$$W_H = (1 - v\chi_H)^{-1} v, \tag{D4}$$

in which v is the bare Coulomb interaction. We use a plane wave cutoff of 8 Ry and we consider 200 bands for the calculation of the cRPA polarization and dielectric function.

2. Choice of the variables in $\mathcal{V}_3$

As a complement to Sec. III C, the choice of the variables in $\mathcal{V}_3$ is discussed below. The DFT band structure is entirely described by the Kohn-Sham energies ϵ_{nk} and orbitals ψ_{nk}. However, it is too complex to consider the whole set $\{\epsilon_{nk}, \psi_{nk}\}$ as the variable space $\mathcal{V}_3$. Instead, we consider variables that capture the essential characteristic energies in the band structure. To choose the variables, we use as a guide the M space. Prioritizing the M space for the choice of variables is natural, for two reasons. First, the M space determines the characteristics of the AB orbital including $|t_1|$ and v [20], because the M space contains the Cu3$d_{x^2-y^2}$-like and O2p_σ-like bands and the AB band. Second, the M space plays a prominent role on the cRPA screening and thus on $R = U/v$. Indeed, the screening increases when the charge-transfer energies between occupied and empty bands decrease [see Eqs. (D1) and (D2)], and the charge-transfer energies are the smallest for the occupied M bands.

3. Values of intermediate quantities within the gMACE procedure

Here, we give the values of intermediate quantities in the variables spaces $\mathcal{V}_s$ with $s = 1 \ldots 4$. On $\mathcal{V}_1$, the values of R_A, R_X, $R_{A'}$, $|Z_X|$, and Z_A that are considered in this paper are represented in Fig. 15. (The numerical values are given in Sec. S1 of the Supplemental Material [21].) The values of R_A, R_X, and $R_{A'}$ are the crystal ionic radii values from Ref. [36]. (For completeness, we mention that ions are assumed to be 6-coordinate in the calculation of these values.) For hole-doped compounds, the partial ion substitution is accounted for as follows: If $A' = \varnothing$, the ion A becomes $A_{2-\delta}\tilde{A}_\delta$ where the atomic number of $\tilde{A}$ is that of A minus one; otherwise, the ion A' is $Hg_{1-\delta}Au_\delta$. In this case, we interpolate linearly the values of ionic radii; namely, we consider

$$R_{A_{2-\delta}\tilde{A}_\delta} = (1 - 0.5\delta)R_A + 0.5\delta R_{\tilde{A}}, \tag{D5}$$

$$R_{Hg_{1-\delta}Au_\delta} = (1 - \delta)R_{Hg} + \delta R_{Au}. \tag{D6}$$

The values of Z_A are interpolated similarly. On $\mathcal{V}_2$, the values of crystal parameters a, c, d_A^z, and d_X^z obtained by the structural optimization are represented in Fig. 16. (These parameters are defined in Table II). We consider $c_\perp = 0$ if $A' = Hg_{1-\delta}Au_\delta$ and $c_\perp = a/2$ if $A' = \varnothing$.

On $\mathcal{V}_3$, the values of variables that are discussed in the main text are represented in Fig. 16. The numerical values of all variables in $\mathcal{V}_3$ and the band structures are given in Secs. S1 and S2 of the Supplemental Material [21], respectively.

APPENDIX E: SCORE ANALYSIS OF THE RESULTS OF THE HIERARCHICAL DEPENDENCE EXTRACTION PROCEDURE

Here, we define the score, and we discuss how the score analysis allows us to construct the physical interpretation of the dependence of y on $\mathcal{V}$ based on the results of the HDE$[y, \mathcal{V}]$. Then, as a complement to Sec. IV, we detail the analysis of the score in the items (I), (III), (IV) and (1)–(10) in Sec. IV.

First, we detail the definition of the score. We define the maximal fitness of the variable x_i at the generation $g = 1$ and $g \geqslant 2$ as, respectively,

$$\tilde{f}_{(1)}[y, x_i] = \max_{\alpha_1} f_{(1)}\left[y, x_i^{\alpha_1}\right], \tag{E1}$$

$$\tilde{f}_{(g)}[y, x_i] = \max_{(\alpha_g,\beta_g,\zeta_g)} f_{(g)}\left[y, x_{(g-1)}^{\text{opt}} \star_{(\zeta_g,\beta_g,\alpha_g)} x_i\right], \tag{E2}$$

which is the maximal value of the fitness function that is obtained at g by enforcing $x_{i_g^{\text{opt}}} = x_i$. Then, we define the score of the variable x_i at the generation g as

$$s_{(g)}[y, x_i] = \frac{\tilde{f}_{(g)}[y, x_i] - \min_{i'} \tilde{f}_{(g)}[y, x_{i'}]}{\max_{i'} \tilde{f}_{(g)}[y, x_{i'}] - \min_{i'} \tilde{f}_{(g)}[y, x_{i'}]}. \tag{E3}$$

We have $s_{(g)}[y, x_i] = 1$ if x_i corresponds to $x_{i_g}^{\text{opt}}$ that is found in the optimization, and $s_{(g)}[y, x_i] = 0$ if x_i has the lowest $\tilde{f}_{(g)}[y, x_i]$ among the variables in $\mathcal{V}$.

Second, we discuss the physical interpretation of the dependence of y on $\mathcal{V}$ and how the score helps to construct such interpretation. The fitness function $f[y, x_{(g)}^{\text{opt}}]$ as defined in Eq. (10) measures the linear correlation between y and $x_{(g)}^{\text{opt}}$. However, the correlation does not always imply a causation, i.e., a physical dependence of y on $x_{(g)}^{\text{opt}}$. Namely, when y and $x_{(g)}^{\text{opt}}$ are correlated, there are two possible scenarios: (i) Causative: There is a physical dependence of y on $x_{i_g^{\text{opt}}}$. (ii) Noncausative: The correlation between y and $x_{i_g^{\text{opt}}}$ is coincidental, and there is no physical dependence of y on $x_{i_g^{\text{opt}}}$.

In practice, we need to judge whether the correlation is causative, especially in the MODg. Below, we discuss how to do so at $g = 1$. Two things should be examined. First,

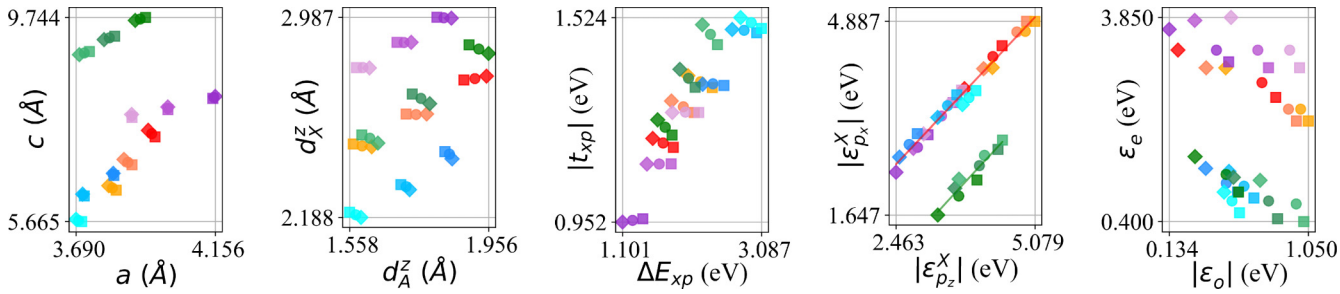

FIG. 16. Values of crystal parameters a, c, d_A^z, and d_X^z obtained by the structural optimization, and values of ΔE_{xp} and $|t_{xp}|$, $|\epsilon^X_{p_z}|$, and $|\epsilon^X_{p_x}|$, and $|\epsilon_o|$ and ϵ_e obtained in the MACE calculation. On $|\epsilon^X_{p_z}|$ and $|\epsilon^X_{p_x}|$, the red and green solid lines correspond to the affine interpolations in Eq. (30) and Eq. (31), respectively. Values of other variables in $\mathcal{V}_3$ are given in Sec. S1 of the Supplemental Material [21]. The CF that corresponds to each color point is shown in Fig. 3.

whether there is a competition between $x_{i_1^{\rm opt}}$ and other variables, namely, whether some variables other than $x_{i_1^{\rm opt}}$ have a score that is close to one. Second, whether the value of the maximal fitness $\tilde{f}[y, x_{i_g^{\rm opt}}]$ [Eq. (E1)] is close to one. There are three possible scenarios (S1), (S2), and (S3):

(S1) *If $x_{i_1^{\rm opt}}$ is not in competition with other variables:* The dominance of $x_{i_1^{\rm opt}}$ in the dependence is unambiguous, and we may identify a physical dependence of y on $x_{i_1^{\rm opt}}$.

(S2) *If competition exists and $\tilde{f}[y, x_{i_1^{\rm opt}}]$ is not close to 1:* then y has a physical dependence on several variables. If y depends mainly on two variables x_i and $x_{i'}$, then each of these variables corresponds to either $x_{i_1^{\rm opt}}$ and $x_{i_2^{\rm opt}}$, and $f[y, x^{\rm opt}_{(2)}]$ has a rather high value, typically $\gtrsim$0.9.

(S3) *If competition exists and $\tilde{f}[y, x_{i_1^{\rm opt}}]$ is close to 1:* The physical dependence is hidden in one of the variables that is highly correlated with y, and for other variables, the correlation with y is coincidental and not physical. Indeed, if $\tilde{f}[y, x_{i_1^{\rm opt}}]$ is close to 1, then y cannot have a strong dependence on more than one variable. If y has a strong dependence on two or more variables, then each of these variables is necessary but not sufficient to describe y accurately. In this case, $\tilde{f}[y, x_{i_1^{\rm opt}}]$ cannot be close to 1, which contradicts (S3). (Typically, we obtain $\tilde{f}[y, x_{i_1^{\rm opt}}] \lesssim 0.7$.)

(I) *Dependence of $|t_1|$ on $\mathcal{V}_1$.* See Fig. 17(a): At $g = 1$, $s_{(1)}[|t_1|, R_{\rm X}] = 1$, but $s_{(1)}[|t_1|, R_{\rm A}]$ is very close to one, and $\tilde{f}_{(1)}[|t_1|, R_{\rm X}] = 0.68$ is not close to one. Then, at $g = 2$ [Eq. (2)], $s_{(2)}[|t_1|, R_{\rm A}] = 1$, $\tilde{f}_{(2)}[|t_1|, R_{\rm A}] = 0.88$, and $R_{\rm A}$ is not in competition with other variables. This corresponds to the scenario (S2), and $R_{\rm X}$ and $R_{\rm A}$ have equal importance in the MOD2 of $|t_1|$. If we reduce the number of compounds in the training set, $(x_{i_1^{\rm opt}}, x_{i_2^{\rm opt}})$ may fluctuate between $(R_{\rm X}, R_{\rm A})$ and $(R_{\rm A}, R_{\rm X})$ as discussed in Appendix F; however, this does not change the physical dependence of $|t_1|$.

(III) *Dependence of v on $\mathcal{V}_1$.* See Fig. 17(b). At $g = 1$, $s_{(1)}[v, R_{\rm A}] = 1$, and $R_{\rm A}$ is not in close competition with other variables. At $g = 2$ [Eq. (4)], $s_{(2)}[v, Z_{\rm A}] = 1$ and $Z_{\rm A}$ is not in close competition with other variables. This corresponds to the scenario (S1). Thus, $R_{\rm A}$ and $Z_{\rm A}$ correspond respectively to $x_{i_1^{\rm opt}}$ and $x_{i_2^{\rm opt}}$ unambiguously. If we reduce the number of compounds in the training set, $x_{i_2^{\rm opt}}$ may occasionally be $n_{\rm AB}$ instead of $Z_{\rm A}$ (see Appendix F); however, this does not change the physical dependence of v.

(IV) *Dependence of R on $\mathcal{V}_1$.* See Fig. 17(c). At $g = 1$, $s_{(1)}[R, |Z_{\rm X}|] = 1$, and $|Z_{\rm X}|$ is not in competition with other variables. At $g = 2$ [Eq. (5)], $s_{(2)}[R, n_{\rm AB}] = 1$ and $n_{\rm AB}$ is not in competition with other variables. This corresponds to the scenario (S1). Thus, $|Z_{\rm X}|$ and $n_{\rm AB}$ correspond respectively to $x_{i_1^{\rm opt}}$ and $x_{i_2^{\rm opt}}$ unambiguously. If we reduce the number of compounds in the training set, $(x_{i_1^{\rm opt}}, x_{i_2^{\rm opt}})$ may occasionally fluctuate (see Appendix F); however, this does not change the physical dependence of R.

(1) *Dependence of $|t_1|$ on $\mathcal{V}_3$.* See Fig. 17(d). At $g = 1$, $s_{(1)}[|t_1|, |t_{xp}|] = 1$, but $s_{(1)}[|t_1|, |t^{O,O'}_{p_\sigma,p_\sigma}|]$, $s_{(1)}[|t_1|, |t^{O,O'}_{p_\sigma,p_\pi}|]$, and $s_{(1)}[|t_1|, W_{\rm M}]$ are close to one, so that $|t_{xp}|$ is in competition with $|t^{O,O'}_{p_\sigma,p_\sigma}|$, $|t^{O,O'}_{p_\sigma,p_\pi}|$, and $W_{\rm M}$. $\tilde{f}_{(1)}[|t_1|, |t_{xp}|] = 0.95$ is rather close to one. This corresponds to the scenario (S3). We identify the physical dependence of $|t_1|$ as that on $|t_{xp}|$. (The intuitive dependence of $|t_1|$ on $|t_{xp}|$ is discussed in Sec. IV.) At $g = 2$ [Eq. (17)], $s_{(2)}[|t_1|, \Delta E_{xp}] = 1$, but $s_{(1)}[|t_1|, W_{\rm M}]$, $s_{(1)}[|t_1|, |\epsilon^{\rm O}_{p_\sigma}|]$, $s_{(1)}[|t_1|, |\epsilon^{\rm O}_{p_\pi}|]$, and $s_{(1)}[|t_1|, |\epsilon^{\rm O}_{p_z}|]$ are also close to one, because ΔE_{xp} is physically correlated with these variables. We identify the physical dependence as that on ΔE_{xp}. (The intuitive dependence of $|t_1|$ on ΔE_{xp} is discussed in Sec. IV.)

(2) *Dependence of $|t_{xp}|$ on $\mathcal{V}_2$.* See Fig. 17(e). At $g = 1$ [Eq. (18)], $s_{(1)}[|t_{xp}|, a] = 1$, and a is not in very close competition with other variables. (We have $s_{(1)}[|t_{xp}|, d^z_{\rm X}] = 0.93$, which is not so far from one but not very close either.) Also, $\tilde{f}_{(1)}[|t_{xp}|, a] = 1.000$. This corresponds to the scenario (S1), and the dominance of a in the MOD1 of $|t_{xp}|$ is unambiguous.

(3) *Dependence of a on $\mathcal{V}_1$.* See Fig. 17(f). At $g = 1$, $s_{(1)}[a, R_{\rm A}] = 1$, $\tilde{f}_{(1)}[a, R_{\rm A}] = 0.80$, and $R_{\rm A}$ is not in close competition with other variables. At $g = 2$ [Eq. (20)], $s_{(2)}[a, R_{\rm X}] = 1$, $\tilde{f}_{(2)}[a, R_{\rm X}] = 0.97$, and $R_{\rm X}$ is not in competition with other variables. This corresponds to the scenario (S1). Thus, $R_{\rm A}$ and $R_{\rm X}$ correspond respectively to $x_{i_1^{\rm opt}}$ and $x_{i_2^{\rm opt}}$ unambiguously.

(4) *Dependence of ΔE_{xp} on $\mathcal{V}_2$.* See Fig. 17(g). At $g = 1$, $s_{(1)}[\Delta E_{xp}, d^z_X] = 1$, but $s_{(1)}[\Delta E_{xp}, a] = 0.97$, so that a is in close competition with d^z_X. We have $\tilde{f}_{(1)}[\Delta E_{xp}, d^z_X] = 0.93$. At $g = 2$ [Eq. (21)], $s_{(2)}[\Delta E_{xp}, a] = 1$ and $\tilde{f}_{(2)}[\Delta E_{xp}, a] = 0.94$, and a is not in close competition with other variables. This corresponds to the scenario (S2): the equal importance of d^z_X and a in the dependence of $|t_{xp}|$ is discussed in the main text. Consistently, in Fig. 7(c), the amplitude of the variation in

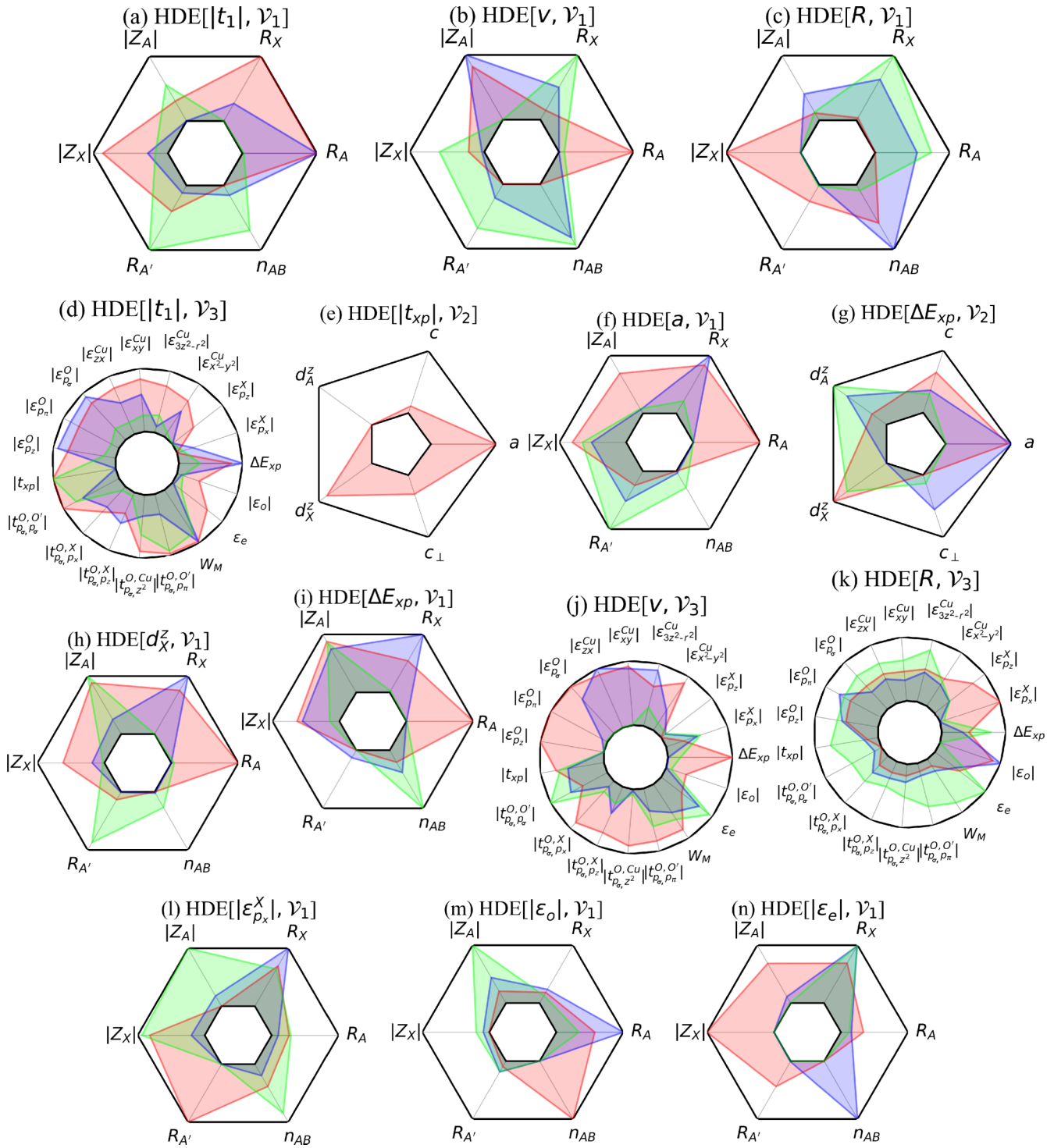

FIG. 17. Score of each variable x_i [Eq. (E3)] in the dependence of y at $g = 1$ (red), $g = 2$ (blue) and $g = 3$ (green) in the HDE[y, $\mathcal{V} = \{x_i\}$], for the items (I), (III), (IV), and (1)–(10) in Sec. IV. The inner and upper bounds of the histogram correspond to $s_{(g)}[y, x_i] = 0$ and $s_{(g)}[y, x_i] = 1$, respectively. The numerical values of the score are given in Sec. S1 of the Supplemental Material [21]. Panel (e) is restricted to $g = 1$.

ΔE_{xp} with either d_X^z or a looks similar: The color map has a diagonal-like pattern.

(5) *Dependence of d_X^z on $\mathcal{V}_1$.* See Fig. 17(h). At $g = 1$, $s_{(1)}[d_X^z, R_\mathrm{A}] = 1$, and R_A is not in close competition with other variables. At $g = 2$ [Eq. (22)], $s_{(2)}[d_X^z, R_\mathrm{X}] = 1$ and R_X is not in competition with other variables. This corresponds to the scenario (S1). Thus, R_A and R_X correspond respectively to $x_{i_1^{\mathrm{opt}}}$ and $x_{i_2^{\mathrm{opt}}}$ unambiguously.

(4, 5) *Dependence of ΔE_{xp} on $\mathcal{V}_1$.* See Fig. 17(i). At $g = 1$, $s_{(1)}[\Delta E_{xp}, R_\mathrm{A}] = 1$, and R_A is not in close competition with

other variables. We have $\tilde{f}_{(1)}[\Delta E_{xp}, R_{\rm A}] = 0.90$. At $g = 2$, $s_{(2)}[\Delta E_{xp}, R_{\rm X}] = 1$, and $R_{\rm X}$ is not in close competition with other variables. We have $\tilde{f}_{(2)}[\Delta E_{xp}, R_{\rm X}] = 0.96$. This corresponds to the scenario (S1). Thus, $R_{\rm A}$ and $R_{\rm X}$ correspond respectively to $x_{i_1^{\rm opt}}$ and $x_{i_2^{\rm opt}}$ unambiguously in Eq. (23).

(6) *Dependence of v on $\mathcal{V}_3$.* See Fig. 17(j). At $g = 1$, $s_{(1)}[v, |\epsilon^{\rm O}_{p_z}|] = 1$, but $s_{(1)}[v, |\epsilon^{\rm O}_{p_\sigma}|]$, $s_{(1)}[v, |\epsilon^{\rm O}_{p_\pi}|]$, and $s_{(1)}[v, \Delta E_{xp}]$ are all close to one, so that $|\epsilon^{\rm O}_{p_z}|$ is in close competition with $|\epsilon^{\rm O}_{p_\sigma}|$, $|\epsilon^{\rm O}_{p_\pi}|$, and ΔE_{xp}. Also, $\tilde{f}_{(1)}[v, |\epsilon^{\rm O}_{p_z}|] = 0.97$ is close to one. This corresponds to the scenario (S3). We identify the physical dependence of v as that on ΔE_{xp} in Sec. IV. At $g = 2$, $s_{(2)}[v, |\epsilon^{\rm Cu}_{z^2}|] = 1$, and $|\epsilon^{\rm Cu}_{z^2}|$ is not in very close competition with other variables. However, the physical meaning of this result is biased by the fact that $|\epsilon^{\rm O}_{p_z}|$ corresponds to $x_{i_1^{\rm opt}}$, which is not physical as discussed above. If we consider the HDE$[v, \{\Delta E_{xp}, |t_{xp}|\}]$, we obtain $s_{(1)}[v, \Delta E_{xp}] = 1$ then $s_{(2)}[v, |t_{xp}|] = 1$ in Eq. (24), as discussed in Sec. IV.

(7) *Dependence of R on $\mathcal{V}_3$.* See Fig. 17(k). At $g = 1$, $s_{(1)}[R, |\epsilon^{\rm X}_{p_x}|] = 1$, and $|\epsilon^{\rm X}_{p_x}|$ is not in close competition with other variables. At $g = 2$, $s_{(2)}[R, |\epsilon_o|] = 1$, and $|\epsilon_o|$ is not in competition with other variables. At $g = 3$ [Eq. (26)], $s_{(3)}[R, \epsilon_e] = 1$ and ϵ_e is not in competition with other variables. This corresponds to the scenario (S1). Thus, $|\epsilon^{\rm X}_{p_x}|$, $|\epsilon_o|$, and ϵ_e correspond respectively to $x_{i_1^{\rm opt}}$, $x_{i_2^{\rm opt}}$, and $x_{i_3^{\rm opt}}$ unambiguously.

(8) *Dependence of $|\epsilon^{\rm X}_{p_x}|$ on $\mathcal{V}_1$.* See Fig. 17(l). At $g = 1$, $s_{(1)}[|\epsilon^{\rm X}_{p_x}|, R_{\rm A'}] = 1$, and $R_{\rm A'}$ is not in close competition with other variables. At $g = 2$ [Eq. (27)], $s_{(2)}[|\epsilon^{\rm X}_{p_x}|, R_{\rm X}] = 1$ and $R_{\rm X}$ is not in competition with other variables. This corresponds to the scenario (S1). Thus, $R_{\rm A'}$ and $R_{\rm X}$ correspond respectively to $x_{i_1^{\rm opt}}$ and $x_{i_2^{\rm opt}}$ unambiguously.

(9) *Dependence of $|\epsilon_o|$ on $\mathcal{V}_1$.* See Fig. 17(m). At $g = 1$, $s_{(1)}[|\epsilon_o|, n_{\rm AB}] = 1$, and $n_{\rm AB}$ is not in competition with other variables. At $g = 2$ [Eq. (28)], $s_{(2)}[|\epsilon_o|, R_{\rm A}] = 1$ and $R_{\rm A}$ is not in competition with other variables. This corresponds to the scenario (S1). Thus, $n_{\rm AB}$ and $R_{\rm A}$ correspond respectively to $x_{i_1^{\rm opt}}$ and $x_{i_2^{\rm opt}}$ unambiguously.

(10) *Dependence of ϵ_e on $\mathcal{V}_1$.* See Fig. 17(n). At $g = 1$, $s_{(1)}[\epsilon_e, |Z_{\rm X}|] = 1$, and $|Z_{\rm X}|$ is not in competition with other variables. At $g = 2$ [Eq. (29)], $s_{(2)}[\epsilon_e, n_{\rm AB}] = 1$ and $n_{\rm AB}$ is not in competition with other variables. This corresponds to the scenario (S1). Thus, $|Z_{\rm X}|$ and $n_{\rm AB}$ correspond respectively to $x_{i_1^{\rm opt}}$ and $x_{i_2^{\rm opt}}$ unambiguously.

APPENDIX F: ROBUSTNESS OF THE PHYSICAL DEPENDENCIES AND MATHEMATICAL EXPRESSIONS OF $|t_1|$ AND u

Here, we examine the robustness of the expressions of $y = |t_1|$, u, v, and R in Eqs. (2), (3), (4), and (5). Namely, we examine how the expressions are modified if the number N_t of compounds in the training set is changed. Also, we show that the physical interpretations of the CF dependencies of $|t_1|$, u, v, and R are robust and independent of N_t.

We define the following procedure, denoted HDE$_{\rm ex}[y, \mathcal{V}]$. We define a number $N_r = 100$ of reduced training sets; each one is a copy of the initial training set of $N_t = 36$ compounds, from which we exclude $N_{t,\rm ex} = 5$ randomly selected compounds. Then, for each reduced training set r, we perform the HDE$[y, \mathcal{V}]$ calculation; the optimized variational parameters are denoted as $\mathbf{p}_g^{{\rm opt,ex},r} = (i_g^{{\rm opt,ex},r}, \alpha_g^{{\rm opt,ex},r}, \zeta_g^{{\rm opt,ex},r}, \beta_g^{{\rm opt,ex},r})$. We obtain N_r different values of $\mathbf{p}_g^{{\rm opt,ex},r}$. Then, we examine how the values of $\mathbf{p}_g^{{\rm opt,ex},r}$ change with respect to $\mathbf{p}_g^{\rm opt}$ that was obtained for the training set of N_t compounds. We restrict the discussion to $g = 1, 2$ and the MOD2 of $y = |t_1|, u, v, R$ on $\mathcal{V}_1$.

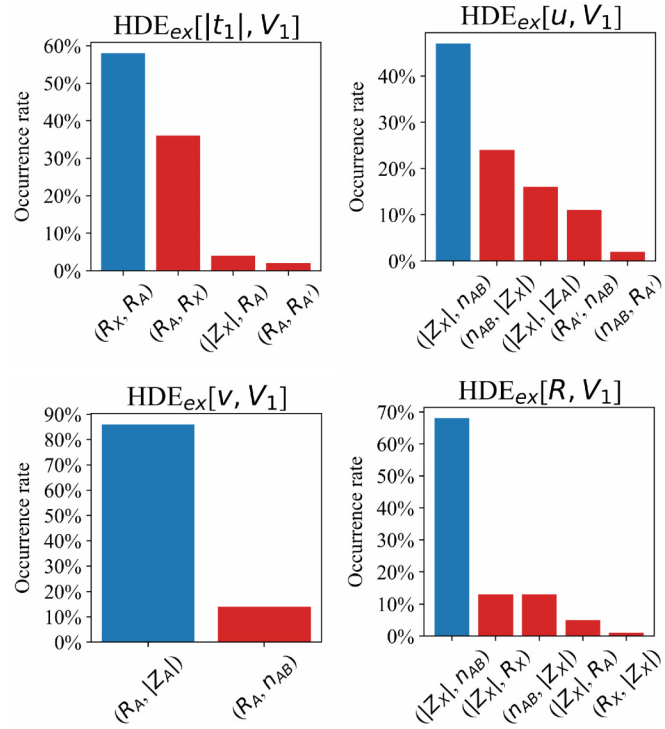

FIG. 18. Occurrence rate of the values of (i_1, i_2) obtained in the $N_r = 100$ HDE$_{\rm ex}[y, \mathcal{V}_1]$ calculations, for $y = |t_1|, u, v, R$. The blue bar corresponds to $(i_1, i_2) = (i_1^{\rm opt}, i_2^{\rm opt})$ obtained in the HDE$[y, \mathcal{V}_1]$ calculation, and the red bars show values of (i_1, i_2) other than $(i_1^{\rm opt}, i_2^{\rm opt})$.

First, we examine the values of $(i_1^{{\rm opt,ex},r}, i_2^{{\rm opt,ex},r})$ obtained in the HDE$_{\rm ex}[y, \mathcal{V}_1]$. In Fig. 18, we show the occurrence rate of (i_1, i_2), that is, $N(i_1, i_2)/N_r$ (in%), where $N(i_1, i_2)$ is the number of reduced training sets for which we obtain $(i_1^{{\rm opt,ex},r}, i_2^{{\rm opt,ex},r}) = (i_1, i_2)$. The values $(i_1, i_2) = (i_1^{\rm opt}, i_2^{\rm opt})$ have the highest occurrence rate; this supports the reliability of the values of $(i_1^{\rm opt}, i_2^{\rm opt})$ that were obtained in the main text. However, other values of $(i_1, i_2) \neq (i_1^{\rm opt}, i_2^{\rm opt})$ have a nonzero occurrence rate. In the following, we discuss the case $(i_1, i_2) = (i_1^{\rm opt}, i_2^{\rm opt})$, and also $(i_1, i_2) \neq (i_1^{\rm opt}, i_2^{\rm opt})$ whose occurrence rate is the highest besides $(i_1^{\rm opt}, i_2^{\rm opt})$. The complete list of values of $\mathbf{p}_1^{{\rm opt,ex},r}$ and $\mathbf{p}_2^{{\rm opt,ex},r}$ for the N_r calculations is given in Sec. S1 of the Supplemental Material [21].

At fixed (i_1, i_2), we examine the quantitative variation of the MOD2 due to the differences between $(\alpha_g^{{\rm opt,ex},r}, \zeta_g^{{\rm opt,ex},r}, \beta_g^{{\rm opt,ex},r})$ and $(\alpha_g^{\rm opt}, \zeta_g^{\rm opt}, \beta_g^{\rm opt})$ as follows. For the $N(i_1, i_2)$ calculations in which $(i_1^{{\rm opt,ex},r}, i_2^{{\rm opt,ex},r}) = (i_1, i_2)$, we gather the expressions of the MOD2s.

These are

$$y_{\mathrm{MOD2}}[r] = h\big[i_1,\ \alpha_1^{\mathrm{opt,ex},r},\ i_2,\ \alpha_2^{\mathrm{opt,ex},r},\ \beta_2^{\mathrm{opt,ex},r},\ \zeta_2^{\mathrm{opt,ex},r}\big], \tag{F1}$$

in which

$$h[i_1,\alpha_1,i_2,\alpha_2,\beta_2,\zeta_2] = k_0 + k_1 x_{i_1}^{\alpha_1}\big[1+\zeta_2 x_{i_2}^{\alpha_2}/\big[x_{i_1}^{\alpha_1}\big]^{\beta_2}\big], \tag{F2}$$

in which k_0 and k_1 are determined by the affine regression and are also functions of $\alpha_1, i_2, \alpha_2, \beta_2, \zeta_2$. We discuss the average value

$$y_{\mathrm{avg}} = \frac{1}{N(i_1,i_2)}\sum_{r=1}^{N_r}\delta_{i_1,i_2}[r]\,y_{\mathrm{MOD2}}[r] \tag{F3}$$

[where $\delta_{i_1,i_2}[r] = 1$ if $(i_1^{\mathrm{opt,ex},r}, i_2^{\mathrm{opt,ex},r}) = (i_1, i_2)$ and zero otherwise], and the difference

$$\Delta y = \max_r \delta_{i_1,i_2}[r]\,|y_{\mathrm{MOD2}}[r] - y_{\mathrm{avg}}|. \tag{F4}$$

By using these notations, we have

$$y_{\mathrm{avg}} - \Delta y \leqslant y_{\mathrm{MOD2}}[r] \leqslant y_{\mathrm{avg}} + \Delta y \tag{F5}$$

for all r that verify $(i_1^{\mathrm{opt,ex},r}, i_2^{\mathrm{opt,ex},r}) = (i_1, i_2)$.

In the case $(i_1, i_2) = (i_1^{\mathrm{opt}}, i_2^{\mathrm{opt}})$, the quantitative uncertainty due to nonzero Δy is overall small. The values of y_{avg} and Δy are shown in Fig. 19. Quantitatively, the dependence of y_{avg} is very similar to that in Fig. 1. Also, the amplitude of Δy does not exceed $\simeq$5%–10% of the amplitude of y_{avg}. The only exception is for $\Delta|t_1|$ at $R_{\mathrm{A}} \geqslant 1.80$ Å: $|t_1|$ may vary up to $\simeq$0.10–0.15 eV. The quantitative uncertainty on the MOD2 of $|t_1|$ is more significant at high values of ionic radii, at which Eq. (2) may predict negative values of $|t_1|$ that are unphysical. Still, the prescription proposed in Sec. V consists in reducing the ionic radii to maximize $|t_1|$: At lower values of ionic radii, the uncertainty on the MOD2 of $|t_1|$ is very low as seen in Fig. 19.

In the case $(i_1, i_2) \neq (i_1^{\mathrm{opt}}, i_2^{\mathrm{opt}})$, we discuss the value of (i_1, i_2) that has the highest occurrence rate besides $(i_1^{\mathrm{opt}}, i_2^{\mathrm{opt}})$. We restrict the discussion to y_{avg}, which is shown in Fig. 20.

First, we discuss the case of $|t_1|$. The value $(i_1, i_2) = (R_{\mathrm{A}}, R_{\mathrm{X}})$ occurs with a non-negligible rate of $\simeq$36% (see Fig. 18). This is because R_{A} is in close competition with R_{X} at $g = 1$ in the HDE$[|t_1|, \mathcal{V}_1]$ (see Appendix E). However, the dependence of y_{avg} for $(i_1, i_2) = (R_{\mathrm{A}}, R_{\mathrm{X}})$ and $(i_1, i_2) = (i_1^{\mathrm{opt}}, i_2^{\mathrm{opt}}) = (R_{\mathrm{X}}, R_{\mathrm{A}})$ is very similar quantitatively (see Figs. 20 and 19). Thus, the physical dependence of $|t_1|$ on R_{X} and R_{A} is robust.

Second, we discuss the case of u. The value $(i_1, i_2) = (n_{\mathrm{AB}}, |Z_{\mathrm{X}}|)$ occurs with a non-negligible rate of $\simeq$24%. In this case, u_{avg} has the same quantitative dependence as that for $(i_1, i_2) = (i_1^{\mathrm{opt}}, i_2^{\mathrm{opt}}) = (|Z_{\mathrm{X}}|, n_{\mathrm{AB}})$: The color map of the dependence of u_{avg} on $(n_{\mathrm{AB}}, |Z_{\mathrm{X}}|)$ in Fig. 20 looks symmetric to that on $(|Z_{\mathrm{X}}|, n_{\mathrm{AB}})$ in Fig. 19.

Third, we discuss the case of v. The value $(i_1, i_2) = (R_{\mathrm{A}}, n_{\mathrm{AB}})$ is the only one besides $(i_1, i_2) = (i_1^{\mathrm{opt}}, i_2^{\mathrm{opt}}) = (R_{\mathrm{A}}, |Z_{\mathrm{A}}|)$ and has a low occurrence rate of $\simeq$14%. At fixed R_{A}, v decreases when n_{AB} decreases. This is consistent with the decrease in v with decreasing Z_{A} in Eq. (4), because n_{AB} decreases with decreasing Z_{A} [see Eq. (16)].

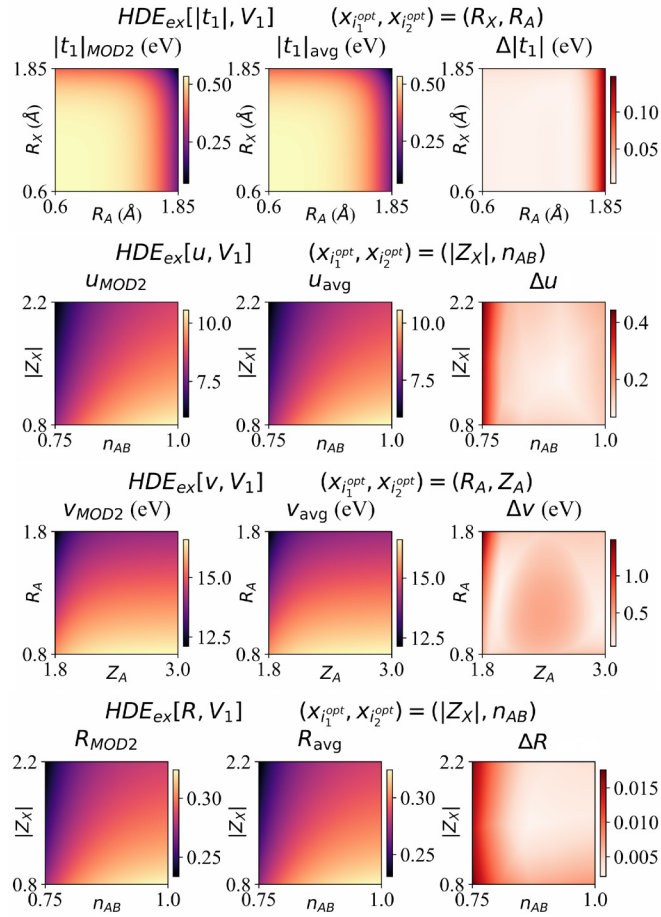

FIG. 19. HDE$_{\mathrm{ex}}[y, \mathcal{V}_1]$ results for $y = |t_1|$, u, v, and R, and for $(i_1, i_2) = (i_1^{\mathrm{opt}}, i_2^{\mathrm{opt}})$. For each y, the left panel shows the MOD2 obtained from the full training set in Fig. 1, the middle panel shows y_{avg} [Eq. (F3)], and the right panel shows Δy [Eq. (F4)].

Fourth, we discuss the case of R. The value $(i_1, i_2) = (n_{\mathrm{AB}}, |Z_{\mathrm{X}}|)$ has an occurrence rate of $\simeq$13%. The color map of the dependence of R on $(n_{\mathrm{AB}}, |Z_{\mathrm{X}}|)$ in Fig. 20 looks symmetric to that on $(|Z_{\mathrm{X}}|, n_{\mathrm{AB}})$ in Fig. 19. The value $(i_1, i_2) = (|Z_{\mathrm{X}}|, R_{\mathrm{X}})$ also has an occurrence rate of $\simeq$13%. At fixed $|Z_{\mathrm{X}}|$, R increases with decreasing R_{X}. This is explained as follows: Decreasing R_{X} decreases the cell parameter a [Eq. (20)], which increases the bandwidth of the M space [20]; this increases the charge-transfer energies between occupied M bands and empty bands, which reduces the screening. [This dependence of R on R_{X} is a higher-order dependence beyond Eq. (5).]

APPENDIX G: HOLE DOPING DEPENDENCE OF SCREENING

Here, as a complement to Sec. IV C and the MOD2 of u and R in Eqs. (3) and (5), we discuss the n_{AB} dependence of u and R when the CF at $\delta = 0$ is fixed. The values of u and R for all compounds in the training set are shown in Fig. 21(a). For X = F, Cl and especially (X, A) = (F, Ba), (Cl, Ba), and (Cl, Sr), we observe a sharp decrease in u between 10% and

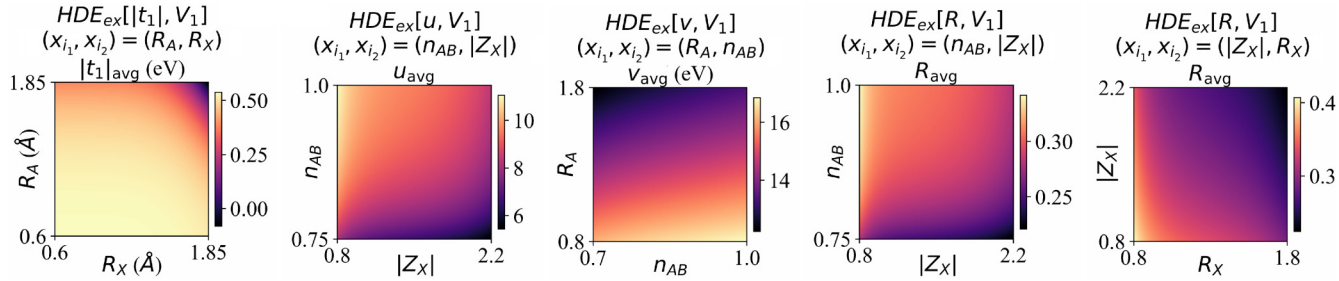

FIG. 20. $HDE_{ex}[y, \mathcal{V}_1]$ results for $y = |t_1|$, u, v, and R, and for selected values of $(i_1, i_2) \neq (i_1^{opt}, i_2^{opt})$ with the highest occurrence rate. For each y, we show y_{avg} [Eq. (F3)].

20% hole doping (i.e., between $n_{AB} = 0.9$ and $n_{AB} = 0.8$). This decrease in u is caused by the sharp decrease in R as seen in Figs. 21(a) and 21(b). Although the decrease in R with increasing δ (or decreasing n_{AB}) is consistent with the MOD2 in Eq. (5), the sharper decrease in R for X = F, Cl compared with X = O is not captured by Eq. (5). For completeness, we discuss the origin of the sharper decrease in R for X = F, Cl in detail here.

The decrease in the δ dependence of R may be quantified by considering the quadratic interpolation

$$R(\delta) = R_0 + R_1\delta + R_2\delta^2, \tag{G1}$$

where R_0, R_1, and R_2 are determined entirely by the *ab initio* values of R at $\delta = 0.0$, 0.1 and 0.2: R_0 is R at $\delta = 0$ in Fig. 21(b), and the values of R_1 and R_2 are given in Fig. 21(c). The decrease in R is encoded in $R_2 < 0$ for all compounds,

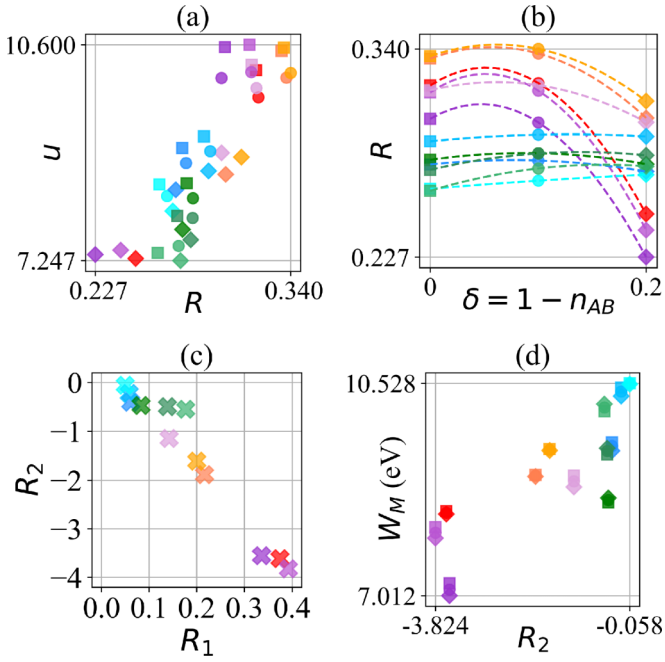

FIG. 21. (a) Values of u and R taken from Fig. 4. (b) Values of R as a function of hole doping $\delta = 1 - n_{AB}$. The dashed curves show the quadratic interpolation of the δ dependence of R by using Eq. (G1). (c) Values of the coefficients R_1 and R_2 in Eq. (G1). (d) Values of W_M as a function of R_2. For each color point, the corresponding CF is shown in Fig. 3. In panel (c), the colors of the crosses correspond to the values of A, X, and A′ in Fig. 3.

and the sharp decrease in R for (X, A) = (F, Ba), (Cl, Ba), and (Cl, Sr) is reflected in the high value of $|R_2|$ compared with the other compounds. Other compounds with X = F, Cl also have higher values of $|R_2|$ compared with X = O. Namely, we have $|R_2| \leqslant 0.55$ for X = O, $|R_2| \geqslant 1.14$ for X = F, Cl, and $|R_2| \geqslant 3.55$ for (X, A) = (F, Ba), (Cl, Ba), and (Cl, Sr).

Possible causes of the higher value of $|R_2|$ are (i) the sharp decrease in $|\epsilon_o|$ from $\delta = 0.1$ to $\delta = 0.2$ for X = F, Cl (see Fig. 16), and also (ii) the lower value of the M space bandwidth W_M for X = F, Cl. On (ii), we see in Fig. 21(d) that the three compounds with (X, A) = (F, Ba), (Cl, Ba), and (Cl, Sr) have the lowest W_M at $\delta = 0.2$. The above discussed dependence of $|R_2|$ on W_M (ii) is interpreted as follows. If W_M is lower, then the charge-transfer energies between the occupied states in the M space and the empty states in the M space (namely, the empty part of the AB band) are smaller. Thus, the intra-M space cRPA screening will be stronger. A rough scaling of the intra-M space screening is $1/W_M$ [see Eq. (D1)], so that the smaller W_M, the more R will decrease when W_M further decreases [and, for X = F, Cl, W_M decreases with increasing δ as seen in Fig. 21(d)]. On the other hand, in X = F, Cl, the screening channel between the occupied states and empty states outside M is relatively weak compared with X = O due to the larger ϵ_e (see Fig. 16). This suggests the intra-M space screening dominates over the other screening channels for X = F, Cl. (Note that, besides the decrease in W_M, the decrease in $|\epsilon_o|$ with increasing δ also participates in increasing the intra-M space screening.)

The higher $|R_2|$ in X = F, Cl compared with X = O implies a nontrivial point on the origin of the superconductivity in oxychlorides and oxyfluorides. [Confirmation of this point requires to improve the derivation of the AB Hamiltonian at the cGW-SIC + LRFB level (this may change the values of R_0, R_1, and R_2), which is left for future studies.] Even though u and R are higher in the undoped compound for X = F, Cl compared with X = O [according to Eq. (3) and (5)], u may become lower in the hole doped compound at optimal hole doping for X = F, Cl compared with X = O due to the higher $|R_2|$. In particular, u may fall into the weak-coupling regime ($u \simeq 6.5$–8.0) at $\delta \simeq 0.2$ for X = F, Cl, which may correspond to the optimal hole doping. For instance, in Table I, the experimental T_c in $Ca_{2-x}Na_xCuO_2Cl_2$ is $T_c^{expt} \simeq 27$ K [12], and this is realized at $\delta = 0.18$, which is close to $\delta = 0.2$ at which the sharp decrease in R happens. If u is in the weak-coupling regime at $\delta = 0.18$, then the lower T_c compared with other cuprates is caused by the decrease in F_{SC} with decreasing u in the weak-coupling regime.

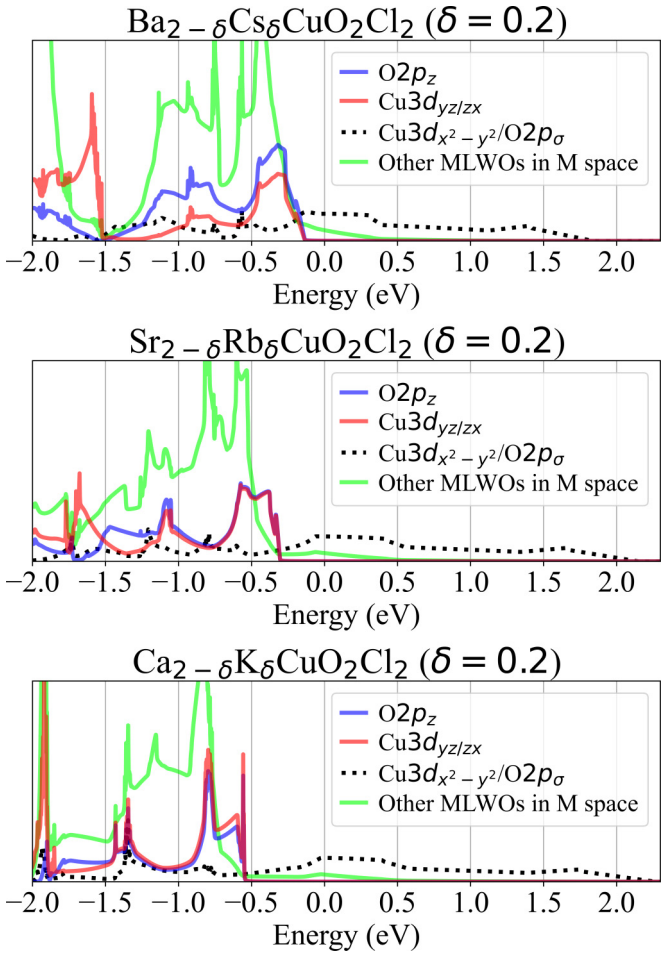

FIG. 22. Partial densities of states within the M space and near the Fermi level for $Ba_{2-\delta}Cs_\delta CuO_2Cl_2$, $Sr_{2-\delta}Rb_\delta CuO_2Cl_2$, and $Ca_{2-\delta}K_\delta CuO_2Cl_2$, at hole doping $\delta = 0.2$.

APPENDIX H: DENSITY OF STATES NEAR THE FERMI LEVEL IN HOLE-DOPED OXYCHLORIDES

In item (9) (dependence of $|\epsilon_o|$ on $\mathcal{V}_1$) in Sec. IV C, we mention the increase in $|\epsilon_o|$ with decreasing R_A in the MOD2 of $|\epsilon_o|$ [Eq. (28)]. Here, we discuss the interpretation of the increase in $|\epsilon_o|$ with decreasing R_A. To do so, we consider as an example the three CFs in the training set such that $\delta = 0.2$, X = Cl and A = $Ba_{2-\delta}Cs_\delta$ (i), $Sr_{2-\delta}Rb_\delta$ (ii), and $Ca_{2-\delta}K_\delta$ (iii). We choose the three above CFs (i)–(iii), because (i) has the lowest value of $|\epsilon_o|$ among the compounds in the training set ($|\epsilon_o| = 0.134$ eV; see the purple diamond marker in Fig. 16), and the value of R_A is progressively reduced from (i) to (iii). [We have $R_A = 1.508$ Å for (i), $R_A = 1.354$ Å for (ii), and $R_A = 1.178$ Å for (iii)].

The interpretation of the increase in $|\epsilon_o|$ with decreasing R_A is based on two observations: (a) The value of $|\epsilon_o|$ mainly depends on the $O2p_z$ and $Cu3d_{yz/zx}$ orbitals, and (b) the $O2p_z$ and $Cu3d_{yz/zx}$ orbitals are deeper in energy when R_A decreases. Below, we discuss items (a) and (b).

First, we discuss item (a). Because ϵ_o is the highest energy of the occupied bands outside the AB band, the value of $|\epsilon_o|$ mainly depends on the orbitals that form the density of states near the Fermi level, excluding the $Cu3d_{x^2-y^2}$ and $O2p_\sigma$ orbitals. Thus, we examine the character of the occupied bands near the Fermi level, for the three above CFs. To do so, we represent the partial density of states (pDOS) for several types of orbitals in Fig. 22. Near the Fermi level, the $Cu3d_{x^2-y^2}/O2p_\sigma$ pDOS corresponds to the AB subspace, and the character of the highest occupied bands outside the AB band is given by the pDOS other than $Cu3d_{x^2-y^2}/O2p_\sigma$. Outside the AB band, the density of states near the Fermi level is dominated by the $O2p_z$ pDOS and $Cu3d_{yz/zx}$ pDOS (see Fig. 22). Thus, the value of $|\epsilon_o|$ mainly depends on the $O2p_z$ and $Cu3d_{yz/zx}$ orbitals. In particular, $|\epsilon_o|$ increases if the $O2p_z$ and $Cu3d_{yz/zx}$ orbitals are deeper in energy.

Now, we discuss (b), i.e., why the $O2p_z$ and $Cu3d_{yz/zx}$ orbitals are deeper in energy when R_A decreases. If R_A is reduced, then a is reduced [see Eq. (20)], and the in-plane O is closer to the A cation. Thus, the positive MP from the A cation that is felt by the in-plane O is stronger. [See Fig. 9(a) for an illustration.] Thus, the $O2p$ orbitals in the M space are stabilized, i.e., their onsite energy is reduced. This shifts the $O2p$ pDOS downward, i.e., farther from the Fermi level. With respect to the O atom in the unit cell, the positions of the four nearest A cations (in Cartesian coordinates) are $\pm a/2\mathbf{y} \pm d_A^z\mathbf{z}$ ($\mathbf{y}$ and $\mathbf{z}$ are unitary vectors along the y and z directions that are considered in Table II): The $O2p_\sigma$ orbital extends along x direction and thus avoids the A cations, whereas the $O2p_z$ orbital extends along z direction, so that the $O2p_z$ electrons are closer to the A cation compared with the $O2p_\sigma$ electrons. This explains why the $O2p_z$ electrons are prominently affected by the positive MP from the A cation. Similarly, the $Cu3d_{yz/zx}$ orbitals extend along the z direction, and may be prominently affected by the positive MP from the A cation. The above discussion is supported by the values of the onsite energies of the orbitals: From (i) to (iii), $\epsilon^O_{p_\sigma}$ decreases by 0.97 eV, whereas $\epsilon^O_{p_z}$ decreases by 1.16 eV, so that the $O2p_z$ orbital is indeed more stabilized than the $O2p_\sigma$ orbital. Also, $\epsilon^{Cu}_{x^2-y^2}$ decreases by 0.26 eV, whereas $\epsilon^{Cu}_{yz/zx}$ decreases by 0.48 eV, so that the $Cu3d_{yz/zx}$ orbital is indeed more stabilized than the $Cu3d_{x^2-y^2}$ orbital.

[1] J. Torrance, Y. Tokura, S. LaPlaca, T. Huang, R. Savoy, and A. Nazzal, New class of high T_c structures: Intergrowth of multiple copper oxide perovskite-like layers with double sheets of BiO, Solid State Commun. **66**, 703 (1988).

[2] L. Gao, Y. Y. Xue, F. Chen, Q. Xiong, R. L. Meng, D. Ramirez, C. W. Chu, J. H. Eggert, and H. K. Mao, Superconductivity up to 164 K in $HgBa_2Ca_{m-1}Cu_mO_{2m+2+\delta}$ ($m = 1$, 2, and 3) under quasihydrostatic pressures, Phys. Rev. B **50**, 4260 (1994).

[3] A. Yamamoto, N. Takeshita, C. Terakura, and Y. Tokura, High pressure effects revisited for the cuprate superconductor family with highest critical temperature, Nat. Commun. **6**, 8990 (2015).

[4] S. N. Putilin, E. V. Antipov, O. Chmaissem, and M. Marezio, Superconductivity at 94 K in $HgBa_2CuO_{4+\delta}$, Nature (London) **362**, 226 (1993).

[5] K. Singh, V. Kirtikar, A. Sinha, and D. Morris, $HgSr_2CuO_{4+\delta}$: A new 78 K superconductor by Mo substitution, Physica C (Amsterdam, Neth.) **231**, 9 (1994).

[6] P. M. Grant, S. S. P. Parkin, V. Y. Lee, E. M. Engler, M. L. Ramirez, J. E. Vazquez, G. Lim, R. D. Jacowitz, and R. L. Greene, Evidence for superconductivity in La_2CuO_4, Phys. Rev. Lett. **58**, 2482 (1987).

[7] J. Attfield, A. Kharlanov, and J. McAllister, Cation effects in doped La_2CuO_4 superconductors, Nature (London) **394**, 157 (1998).

[8] M. AI-Mamouri, P. P. Edwards, C. Greaves, and M. Slaski, Synthesis and superconducting properties of the strontium copper oxy-fluoride $Sr_2CuO_2F_{2+\delta}$, Nature (London) **369**, 382 (1994).

[9] P. Slater, J. Hodges, M. Francesconi, P. Edwards, C. Greaves, I. Gameson, and M. Slaski, An improved route to the synthesis of superconducting copper oxyfluorides $Sr_{2-x}A_xCuO_2F_{2+\delta}$ (A = Ca, Ba) using transition metal difluorides as fluorinating reagents, Physica C (Amsterdam, Neth.) **253**, 16 (1995).

[10] T. Tatsuki, A. Tokiwa-Yamamoto, T. Tamura, Y. Moriwaki, X.-J. Wu, S. Adachi, and K. Tanabe, A new oxychloride, $Ba_2CuO_2Cl_2$, prepared under high pressure, Physica C (Amsterdam, Neth.) **265**, 323 (1996).

[11] Z. Hiroi, N. Kobayashi, and M. Takano, Probable hole-doped superconductivity without apical oxygens in $(Ca, Na)_2CuO_2Cl_2$, Nature (London) **371**, 139 (1994).

[12] K.-H. Kim, H.-J. Kim, J.-D. Kim, H.-G. Lee, and S.-I. Lee, Superconducting properties of $Ca_{1.82}Na_{0.18}CuO_2Cl_2$ single crystals from the equilibrium magnetization, J. Korean Phys. Soc. **48**, 1032 (2006).

[13] Y. Zenitani, S. Sahoda, J. Akimitsu, N. Kubota, and M. Ayabe, New superconductor with apical bromine $(Ca, A)_2CuO_2Br_2$: A = Na, K, J. Jpn. Soc. Powder Powder Metall. **43**, 1087 (1996).

[14] M. T. Schmid, J.-B. Morée, R. Kaneko, Y. Yamaji, and M. Imada, Superconductivity studied by solving *ab initio* low-energy effective Hamiltonians for carrier doped $CaCuO_2$, $Bi_2Sr_2CuO_6$, $Bi_2Sr_2CaCu_2O_8$, and $HgBa_2CuO_4$, Phys. Rev. X **13**, 041036 (2023).

[15] M. Imada and T. Miyake, Electronic structure calculation by first principles for strongly correlated electron systems, J. Phys. Soc. Jpn. **79**, 112001 (2010).

[16] M. Hirayama, T. Miyake, and M. Imada, Derivation of static low-energy effective models by an *ab initio* downfolding method without double counting of Coulomb correlations: Application to $SrVO_3$, FeSe, and FeTe, Phys. Rev. B **87**, 195144 (2013).

[17] M. Hirayama, Y. Yamaji, T. Misawa, and M. Imada, *Ab initio* effective Hamiltonians for cuprate superconductors, Phys. Rev. B **98**, 134501 (2018).

[18] M. Hirayama, T. Misawa, T. Ohgoe, Y. Yamaji, and M. Imada, Effective Hamiltonian for cuprate superconductors derived from multiscale *ab initio* scheme with level renormalization, Phys. Rev. B **99**, 245155 (2019).

[19] J.-B. Morée, M. Hirayama, M. T. Schmid, Y. Yamaji, and M. Imada, *Ab initio* low-energy effective Hamiltonians for the high-temperature superconducting cuprates $Bi_2Sr_2CuO_6$, $Bi_2Sr_2CaCu_2O_8$, $HgBa_2CuO_4$, and $CaCuO_2$, Phys. Rev. B **106**, 235150 (2022).

[20] J.-B. Morée, Y. Yamaji, and M. Imada, Dome structure in pressure dependence of superconducting transition temperature for $HgBa_2Ca_2Cu_3O_8$: Studies by *ab initio* low-energy effective Hamiltonian, Phys. Rev. Res. **6**, 023163 (2024).

[21] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevB.110.014502 for the values of physical quantities that are considered in the gMACE procedure and the band structures for all compounds in the training set, the full list of optimized variational parameters in the HDE and HDE_{ex} calculations, and details on the numerical aspects of HDE.

[22] M. Hirayama, T. Misawa, T. Miyake, and M. Imada, *Ab initio* studies of magnetism in the iron chalcogenides FeTe and FeSe, J. Phys. Soc. Jpn. **84**, 093703 (2015).

[23] M. Hirayama, T. Miyake, M. Imada, and S. Biermann, Low-energy effective Hamiltonians for correlated electron systems beyond density functional theory, Phys. Rev. B **96**, 075102 (2017).

[24] M. Hirayama, M. T. Schmid, T. Tadano, T. Misawa, and M. Imada, *Ab initio* material design of Ag-based oxides for high-T_c superconductor, arXiv:2207.12595.

[25] J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized gradient approximation made simple, Phys. Rev. Lett. **77**, 3865 (1996).

[26] F. Aryasetiawan, M. Imada, A. Georges, G. Kotliar, S. Biermann, and A. I. Lichtenstein, Frequency-dependent local interactions and low-energy effective models from electronic structure calculations, Phys. Rev. B **70**, 195104 (2004).

[27] F. Aryasetiawan, K. Karlsson, O. Jepsen, and U. Schönberger, Calculations of Hubbard U from first-principles, Phys. Rev. B **74**, 125106 (2006).

[28] N. Marzari and D. Vanderbilt, Maximally localized generalized Wannier functions for composite energy bands, Phys. Rev. B **56**, 12847 (1997).

[29] I. Souza, N. Marzari, and D. Vanderbilt, Maximally localized Wannier functions for entangled energy bands, Phys. Rev. B **65**, 035109 (2001).

[30] T. Miyake, F. Aryasetiawan, and M. Imada, *Ab initio* procedure for constructing effective models of correlated materials with entangled band structure, Phys. Rev. B **80**, 155134 (2009).

[31] L. Nordheim, The electron theory of metals, Ann. Phys. (Berlin, Ger.) **401**, 607 (1931).

[32] B. Grande and H. Müller-Buschbaum, Über Oxocuprate. XVII. Zur Kenntnis von $Ca_2CuO_2Cl_2$ und $Ca_2CuO_2Br_2$, Z. Anorg. Allg. Chem. **429**, 88 (1977).

[33] D. N. Argyriou, J. D. Jorgensen, R. L. Hitterman, Z. Hiroi, N. Kobayashi, and M. Takano, Structure and superconductivity without apical oxygens in $(Ca,Na)_2CuO_2Cl_2$, Phys. Rev. B **51**, 8434 (1995).

[34] B. Grande and H. Müller-Buschbaum, Über Erdalkalimetalloxocuprate, VIII Zur Kenntnis von $Sr_2CuO_2Cl_2$, Z. Anorg. Allg. Chem. **417**, 68 (1975).

[35] L. L. Miller, X. L. Wang, S. X. Wang, C. Stassis, D. C. Johnston, J. Faber, and C.-K. Loong, Synthesis, structure, and properties of $Sr_2CuO_2Cl_2$, Phys. Rev. B **41**, 1921 (1990).

[36] R. D. Shannon, Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides, Acta Crystallogr. Sect. A: Cryst. Phys. Diffr. Theor. Gen. Crystallogr. **32**, 751 (1976).

[37] J. D. Jorgensen, H. B. Schüttler, D. G. Hinks, D. W. Capone II, K. Zhang, M. B. Brodsky, and D. J. Scalapino, Lattice instability and high-T_c superconductivity in $La_{2-x}Ba_xCuO_4$, Phys. Rev. Lett. **58**, 1024 (1987).

[38] D. Li, K. Lee, B. Y. Wang, M. Osada, S. Crossley, H. R. Lee, Y. Cui, Y. Hikita, and H. Y. Hwang, Superconductivity in an infinite-layer nickelate, Nature (London) **572**, 624 (2019).

[39] M. Kitatani, L. Si, P. Worm, J. M. Tomczak, R. Arita, and K. Held, Optimizing superconductivity: From cuprates via nickelates to palladates, Phys. Rev. Lett. **130**, 166002 (2023).

[40] K. Momma and F. Izumi, *VESTA3* for three-dimensional visualization of crystal, volumetric and morphology data, J. Appl. Crystallogr. **44**, 1272 (2011).

[41] C. Kim, P. J. White, Z.-X. Shen, T. Tohyama, Y. Shibata, S. Maekawa, B. O. Wells, Y. J. Kim, R. J. Birgeneau, and M. A. Kastner, Systematics of the photoemission spectral function of cuprates: Insulators and hole- and electron-doped superconductors, Phys. Rev. Lett. **80**, 4245 (1998).

[42] I. A. Nekrasov, N. S. Pavlov, E. Z. Kuchinskii, M. V. Sadovskii, Z. V. Pchelkina, V. B. Zabolotnyy, J. Geck, B. Büchner, S. V. Borisenko, D. S. Inosov, A. A. Kordyuk, M. Lambacher, and A. Erb, Electronic structure of $Pr_{2-x}Ce_xCuO_4$ studied via ARPES and LDA + DMFT + $\Sigma_{\mathbf{k}}$, Phys. Rev. B **80**, 140510(R) (2009).

[43] M. Ikeda, T. Yoshida, A. Fujimori, M. Kubota, K. Ono, H. Das, T. Saha-Dasgupta, K. Unozawa, Y. Kaga, T. Sasagawa, and H. Takagi, Effects of chemical pressure on the fermi surface and band dispersion of the electron-doped high-T_c superconductors, Phys. Rev. B **80**, 014510 (2009).

[44] M. Ikeda, T. Yoshida, A. Fujimori, M. Kubota, K. Ono, Y. Kaga, T. Sasagawa, and H. Takagi, Differences in the high-energy kink between hole- and electron-doped high-T_c superconductors, Phys. Rev. B **80**, 184506 (2009).

[45] J. W. Harter, L. Maritato, D. E. Shai, E. J. Monkman, Y. Nie, D. G. Schlom, and K. M. Shen, Nodeless superconducting phase arising from a strong (π, π) antiferromagnetic phase in the infinite-layer electron-doped $Sr_{1-x}La_xCuO_2$ compound, Phys. Rev. Lett. **109**, 267001 (2012).

[46] J. W. Harter, L. Maritato, D. E. Shai, E. J. Monkman, Y. Nie, D. G. Schlom, and K. M. Shen, Doping evolution and polar surface reconstruction of the infinite-layer cuprate $Sr_{1-x}La_xCuO_2$, Phys. Rev. B **92**, 035149 (2015).

[47] K. Tanaka, T. Yoshida, A. Fujimori, D. H. Lu, Z.-X. Shen, X.-J. Zhou, H. Eisaki, Z. Hussain, S. Uchida, Y. Aiura, K. Ono, T. Sugaya, T. Mizuno, and I. Terasaki, Effects of next-nearest-neighbor hopping t' on the electronic structure of cuprate superconductors, Phys. Rev. B **70**, 092503 (2004).

[48] C. T. Shih, T. K. Lee, R. Eder, C.-Y. Mou, and Y. C. Chen, Enhancement of pairing correlation by t' in the two-dimensional extended $t - J$ model, Phys. Rev. Lett. **92**, 227002 (2004).

[49] E. Pavarini, I. Dasgupta, T. Saha-Dasgupta, O. Jepsen, and O. K. Andersen, Band-structure trend in hole-doped cuprates and correlation with T_{cmax}, Phys. Rev. Lett. **87**, 047003 (2001).

[50] S. Tan, Y. Liu, Y. Mou, and S. Feng, Anisotropic dressing of electrons in electron-doped cuprate superconductors, Phys. Rev. B **103**, 014503 (2021).

[51] T. Ohgoe, M. Hirayama, T. Misawa, K. Ido, Y. Yamaji, and M. Imada, *Ab initio* study of superconductivity and inhomogeneity in a Hg-based cuprate superconductor, Phys. Rev. B **101**, 045124 (2020).

[52] Y. Arao, M. Tange, M. Yokoshima, H. Ikeda, and R. Yoshizaki, Optimization of the Bi-2201 superconductors with Pb and La co-doping, Physica C (Amsterdam, Neth.) **426-431**, 351 (2005).

[53] K. Pearson, LIII. On lines and planes of closest fit to systems of points in space, Philos. Mag. **2**, 559 (1901).

[54] K. Ido, Y. Motoyama, K. Yoshimi, and T. Misawa, Data analysis of ab initio effective Hamiltonians in iron-based superconductors — Construction of predictors for superconducting critical temperature, J. Phys. Soc. Jpn. **92**, 064702 (2023).

[55] C. M. Bishop, *Pattern Recognition and Machine Learning* (Springer, Berlin, 2006).

[56] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos *et al.*, QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials, J. Phys.: Condens. Matter **21**, 395502 (19pp) (2009).

[57] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A. D. Corso, S. de Gironcoli, P. Delugas, R. A. DiStasio Jr, A. Ferretti, A. Floris, G. Fratesi, G. Fugallo *et al.*, Advanced capabilities for materials modelling with QUANTUM ESPRESSO, J. Phys.: Condens. Matter **29**, 465901 (2017).

[58] D. R. Hamann, Optimized norm-conserving Vanderbilt pseudopotentials, Phys. Rev. B **88**, 085117 (2013).

[59] M. Schlipf and F. Gygi, Optimization algorithm for the generation of ONCV pseudopotentials, Comput. Phys. Commun. **196**, 36 (2015).

[60] V. Petricek, Y. Gao, P. Lee, and P. Coppens, X-ray analysis of the incommensurate modulation in the 2 : 2 : 1 : 2 Bi-Sr-Ca-Cu-O superconductor including the oxygen atoms, Phys. Rev. B **42**, 387 (1990).

[61] V. Shamray, A. Mikhailova, and A. Mitin, Crystal structure and superconductivity of Bi-2223, Crystallogr. Rep. **54**, 584 (2009).

[62] K. Nakamura, Y. Yoshimoto, Y. Nomura, T. Tadano, M. Kawamura, T. Kosugi, K. Yoshimi, T. Misawa, and Y. Motoyama, RESPACK: An *ab initio* tool for derivation of effective low-energy model of material, Comput. Phys. Commun. **261**, 107781 (2021).