Theoretical proposal and material realization of ferromagnetic negative charge-transfer energy insulator

Zhao Liu,[1] Xingxing Li,[2] W. Zhu,[3] Z. F. Wang,[4] and Jinlong Yang[2,*]

[1]*Hefei National Laboratory for Physical Sciences at the Microscale, University of Science and Technology of China, Hefei, Anhui 230026, China*

[2]*Hefei National Laboratory for Physical Sciences at the Microscale, Synergetic Innovation Center of Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei, Anhui 230026, China*

[3]*Institute of Natural Sciences, Westlake Institution of Advanced Study and School of Science, Westlake University, Hangzhou 310024, China*

[4]*Hefei National Laboratory for Physical Sciences at the Microscale, CAS Key Laboratory of Strongly-Coupled Quantum Matter Physics, University of Science and Technology of China, Hefei, Anhui 230026, China*

(Received 8 August 2022; revised 21 November 2022; accepted 9 January 2023; published 18 January 2023)

Here we propose a type of ferromagnetic semiconductors: ferromagnetic negative charge-transfer energy insulator (FNCTEI). In FNCTEI, the negative charge-transfer states strongly enhance the ferromagnetic (FM) exchange interactions and the orbital hybridization gap permits the magnetic molecular orbitals as the underlying magnetic units rather than local atomic orbitals. Thus, the FM exchange interactions are rather strong and decay slowly due to the large spreading of magnetic molecular orbitals. This is distinct from the superexchange mechanism where FM exchange interactions are quite weak as summarized in the well-known Goodenough-Kanamori-Anderson semiempirical rules. Through first-principle calculations with the hybrid functional, PbO-type CrAs monolayer is mapped out to be a FNCTEI, which possesses a band gap ∼0.35 eV, FM nearest-/next-nearest-neighbor exchange coupling strength ∼57/40 meV, and a high $T_c \sim 1500$ K respectively. It is believed that the theoretical prediction of FNCTEI supports the hypothesis by Khomskii and Sawatzky in 1997 [Solid State Commun. **102**, 87 (1997)].

DOI: 10.1103/PhysRevB.107.014413

I. INTRODUCTION

Ferromagnetism (FM), one of the oldest but most mysterious phenomena, is still intriguing intensive studies [1–5]. The existence/vanishing of charge gap with FM gives itinerant FM/FM semiconductor (or insulator). Different mechanisms have been proposed for itinerant FM, like the early Nagaoka's theorem [6,7], flat-band ferromagnetism [8–13], multiorbital system [14,15], etc. For ferromagnetic semiconductor (FMSC), the underlying mechanism varies for intrinsic and extrinsic system. The extrinsic system, also known as diluted FMSC, is obtained via doping magnetic ions into the nonmagnetic intrinsic semiconductor [16,17]. The multiple degree of freedoms, including charge, spin, orbital, and impurity, make it hard to write down a unified theory [4,18–20]. At the same time, the difficulty in manipulating magnetic impurities greatly hinders their developments [21]. Herein we will focus on intrinsic FMSC.

The intrinsic combination between FM and gap can trace back to the idea of superexchange interaction, first proposed by Kramers [22] and then developed by Anderson [23]. In contrast with the itinerant FM where direct exchange between correlated orbitals are possible, superexchange interaction relies on ligand p orbitals to mediate long-range exchange interactions. Later on several quantitative relations on FM and AFM superexchange interaction were unveiled, mainly by Goodenough [24,25], Kanamori [26], and Anderson [27].

In the 1990s, Zaanen, Sawatzky, and Allen solved the Anderson impurity model [28,29] and classified all gapped transition metal compounds into two types: charge-transfer insulator (CTI) and Mott-Hubbard insulator (MHI) (see Fig. 1). These two insulators are indistinguishable at the ground-state level (labelled as $|d^n\rangle$), but they have different low-energy excitations. Among the two characteristic one-particle excitations in transition metal compounds: $\Delta_{\mathrm{CT}} = E(|d^{n+1}\underline{L}\rangle) - E(|d^n\rangle)$ ($\underline{L}$ means a hole on L), which describes charge transfer from p to d and $U_d = E(|d^{n+1}d^{n-1}\rangle) - E(|d^n d^n\rangle)$, which characterizes charge fluctuation inner d shell, we have $\Delta_{\mathrm{CT}} < U_d$ in CTI while $\Delta_{\mathrm{CT}} > U_d$ in MHI. These low-energy excitations give rise to exchange interaction (with strength labeled by J) when the fermionic model is reduced to spin model. Since the condition $t^{dp} \ll \{\Delta_{\mathrm{CT}}, U_d\}$ (where t^{dp} is the hopping strength between d and p orbitals) is satisfied in both CTI and MHI, a perturbative treatment of kinetic energy is possible. Under such circumstances, the exchange interactions can be derived in a unified form for both CTI and MHI. The obtained semiempirical rules thus cover the aforementioned quantitative relations by Goodenough [24,25], Kanamori [26], and Anderson [27] and are summarized as Goodenough-Kanamori-Anderson (GKA) semiempirical rules nowadays [30], which lays the foundation of modern-day understanding of superexchange interaction. However GKA semiempirical rules impose strong constraints on FM superexchange

*jlyang@ustc.edu.cn

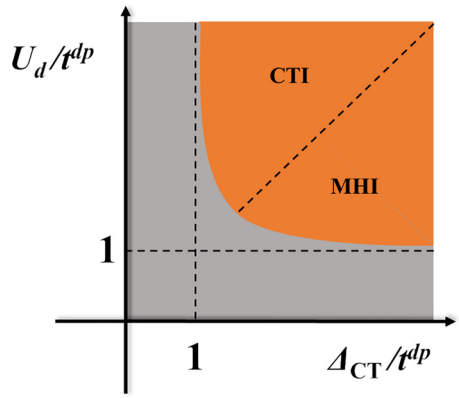

FIG. 1. Schematic illustration of Zaanen-Sawatzky-Allen scheme. The grey (orange) region is metallic (insulating) phase.

coupling strength in both CTI and MHI: The antiferromagnetic (AFM) is generally much stronger than FM superexchange interaction and FMSCs seldom have ambient Curie temperature (T_c), like the long-known bulk EuS (17 K) [31], EuO (69 K) [32], and recently discovered 2D CrI_3 (45 K) [33], $Cr_2Ge_2Te_6$ (30 K) [34]. Deviations from the standard 90° and 180° $d-p-d$ geometries are often seen in transition metal compounds. If the deviation is small, the above GKA semiempirical rules are still available. When the deviation is large, more and more exchange channels are possible, GKA semiempirical rules are less predictive. In this situation, first-principles calculation is a powerful tool to determine the competition. For materials with irregular $d-p-d$ angles between 90° and 180°, the AFM (FM) J is weakened (strengthened) with reducing $d-p-d$ angle; therefore, it is possible to find a balance point with strong FM [35].

As room-temperature FMSCs are the core unit of next-generation spintronic devices, such as processing-in-memory, spin field-effect transistors, magnetic tunneling junctions, and so on [36–38], an alternative mechanism, which can break the inborn bottleneck in the superexchange mechanism is thus highly needed. Nevertheless, it seems that there is no other insulator phase besides MHI or CTI in Fig. 1. Fortunately, this is not the case and in recent years, a hidden insulating phase is discovered in the previously believed metallic phase of Fig. 1 and is now known as negative CTI [39–42]. Since the charge gap is negative, holes are self-doped to the ligand p orbitals even at the ground-state level [39]. It is believed that there is strong AFM direct exchange interaction between metal cations and ligand p holes, therefore Khomskii and Sawatzky guessed a strong emergent FM J between two metal cations in such kind of insulators in 1997 [43] [see Fig. 5(a)]. Although been suggested for a long time, less improvements have been made along this line [44] when comparing with the well-established super-exchange mechanism.

In this paper, we give a systematic investigation on this idea, which motivate us to propose the concept of ferromagnetic negative charge-transfer energy insulator (FNCTEI). This is a type of FMSC with physical properties different from that of FM CTI or MHI. This paper is organized as follows: in Sec. II, we set the stage by introducing the multiband Hamiltonian to describe transition metal compounds. After a perturbative treatment on 90° and 180° $d-p-d$ cluster models and a mean-field treatment on the bipartite square lattice, the basic features of FNCTEI are sketched. In Sec. III, we focus on the the material realization of FNCTEI. In principle, FNCTEI can be realized in any lattice structure by varying related parameters. Here we take PbO-type monolayer as a prototype. By choosing Cr (As) as the transition metal (ligand), CrAs monolayer is found to be an ideal FNCTEI. In the closure Sec. IV, a detailed comparison between FNCTEI and FM CTI (MHI) is made. What is more, other aspects of FNCTEI are also discussed there.

II. HAMILTONIAN FOR TRANSITION METAL COMPOUNDS

Transition metal compounds are described by the multiband $d-p$ model [45]

$$\hat{H} = \hat{H}_d + \hat{H}_p + \hat{H}_{dp}, \tag{1}$$

where $\hat{H}_d$ ($\hat{H}_p$) describes d (p) shell of metal cations (ligand anions) and $\hat{H}_{dp}$ is the intershell term between d and p shells. For simplicity, we will use "M" and "L" for metal and ligand from now on. Each component in Eq (1) is contributed by kinetic and interaction terms. For example, $\hat{H}_d$ is given by

$$\begin{aligned}\hat{H}_d = {} & \varepsilon_d \sum_{i,m,\sigma} \hat{n}^{(d)}_{im\sigma} + \sum_{i,j,m,n,\sigma} \left(t^d_{im,jn} \hat{d}^\dagger_{im\sigma} \hat{d}_{jn\sigma} + \text{H.c.}\right) \\ & + u_d \sum_{i,m} \sum_{i,m} \hat{n}^{(d)}_{im\uparrow} \hat{n}^{(d)}_{im\downarrow} + u'_d \sum_{i,m\neq n} \hat{n}^{(d)}_{im\uparrow} \hat{n}^{(d)}_{in\downarrow} \\ & + \frac{u'_d - j^{(d)}_H}{2} \sum_{m\neq n,\sigma} \hat{n}^{(d)}_{im\sigma} \hat{n}^{(d)}_{in\sigma} \\ & - j^{(d)}_H \sum_{m\neq n} (\hat{d}^\dagger_{im\uparrow} \hat{d}_{im\downarrow} \hat{d}^\dagger_{in\downarrow} \hat{d}_{in\uparrow} - \hat{d}^\dagger_{im\uparrow} \hat{d}^\dagger_{im\downarrow} \hat{d}_{in\downarrow} \hat{d}_{in\uparrow}),\end{aligned} \tag{2}$$

where i, m, σ are indices for the site, orbital, and spin degree of freedom, $\hat{d}^\dagger_{im\sigma}$ ($\hat{d}_{im\sigma}$) is the creation (annihilation) operator for a d electron labeled by site i, orbital m and spin σ. The first term is the on-site energy (ε_d) of d orbitals. The second term is the hopping energy (t^d) between different d orbitals at different site, which gives d bands. The left term describes the intra-atomic Coulomb interactions expressed by Kanamori parameters u_d, u'_d and $j^{(d)}_H$. To retain the rotational invariance in real space, we have the constraint $u'_d = u_d - 2j^{(d)}_H$.

As for p shell, $\hat{H}_p$ have a similar form as $\hat{H}_d$, and the corresponding creation (annihilation) operator is $\hat{p}^\dagger_{im\sigma}$ ($\hat{p}_{im\sigma}$), with the parameters ε_p, t^p, u_p, u'_p, and $j^{(p)}_H$. These intra-atomic Coulomb interactions are usually ignored, but in fact, these terms are not small at all, especially the Hund's coupling for p (O: 1.2 eV) can be even larger than d orbitals.

The last component is generally taken as hopping energy and intershell density-density interaction

$$\begin{aligned}\hat{H}_{dp} = {} & \sum_{i,j,m,n,\sigma} \left(t^{dp}_{im,jn} \hat{d}^\dagger_{im\sigma} \hat{p}_{jn\sigma} + \text{H.c}\right) \\ & + u_{dp} \sum_{\langle i,j\rangle,m,n,\sigma,\sigma'} \hat{n}^{(d)}_{im\sigma} \hat{n}^{(p)}_{jn\sigma'},\end{aligned} \tag{3}$$

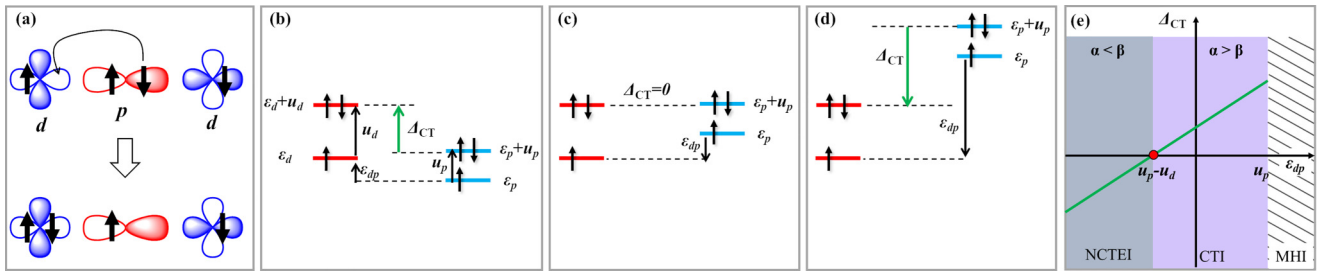

FIG. 2. (a) Real space description of Δ_{CT} in a cluster model formed by a p connecting two d orbitals. [(b)–(d)] Energy space description of Δ_{CT} for $u_d > \Delta_{CT} > 0$, $\Delta_{CT} = 0$, and $\Delta_{CT} < 0$ for model in (a). Here, only ε_p is the variable with all other parameters fixed. (e) A schematic phase diagram for Δ_{CT} vs ε_{dp}.

where t^{dp} is the orbital hybridization between d and p shell, which is responsible for crystal field splitting and u_{dp} describes the nonlocal intershell Coulomb interaction. Believed to be small, u_{dp} is always abandoned so only $p-d$ hopping are considered. However, such an interaction is important in stabilizing charge gap in cuprates [46,47] and modifying exchange interactions [48].

The large parameter space makes the exact solution of Eq. (1) intractable and approximations are necessary. For example, p shell always lies below d shell, thus it is reasonable to treat d shell as the only active shell. Under such circumstance, it is reasonable to ignore $\hat{H}_p$, $\hat{H}_{dp}$ in Eq. (1) and the full model is simplified to the Anderson lattice model (ALM) with screened u_d, u'_d, and $j_H^{(d)}$. What is more, if the d bands are narrow, it is safe to ignore t^d and ALM is further reduced to Anderson impurity model, which has been solved by Zaanen *et al.* as mentioned above [28,29]. However, when p and d orbitals are both active, the above assumption fails as both p and d orbitals should be treated as correlated.

A. Negative charge-transfer energy insulator

Since Δ_{CT} describes an electron hopping from p to d orbitals, it is a function of ε_d, ε_p and the interaction parameters,

$$\Delta_{CT} = \varepsilon_{dp} + E_{int}\left(u_d, j_H^{(d)}, u_p, j_H^{(p)}\right), \tag{4}$$

where ε_{dp} is the on-site energy difference between d and p orbitals: $\varepsilon_{dp} = \varepsilon_d - \varepsilon_p$ and $E_{int}(u_d, j_H^{(d)}, u_p, j_H^{(p)})$ describes the interaction contribution to Δ_{CT}. A general expression of $E_{int}(u_d, j_H^{(d)}, u_p, j_H^{(p)})$ is impossible as it depends on the d and p fillings. For the electron hopping process represented in Fig. 2(a), we have $E_{int}(u_d, j_H^{(d)}, u_p, j_H^{(p)}) = u_d - u_p$ as p orbital is no longer fully occupied and d orbital becomes fully occupied [seen in Fig. 2(b)] so

$$\Delta_{CT} = \varepsilon_{dp} + u_d - u_p. \tag{5}$$

In most transition metal compounds, p orbitals are below d orbitals and if $\varepsilon_{dp} < u_p$, the system is a CHI as depicted in Fig. 2(b). Suppose we can tune ε_p to higher energy (with ε_d fixed), a special point is that $\varepsilon_{dp} = u_p - u_d$ (or $\Delta_{CT} = 0$) as shown in Fig. 2(c). At this point, since the transfer of the electron from p to d orbital does not consume energy, state $|d^1\rangle$ and $|d^2\underline{L}\rangle$ have the same energy and the system is gapless. If ε_p continues to increase, then Δ_{CT} becomes negative and the system becomes negative charge-transfer energy insulator [NCTEI, see Fig. 2(d)]. Configuration interaction tells us the ground state should be a linear combination of $|d^1\rangle$ and $|d^2\underline{L}\rangle$ when t^{dp} is included,

$$|GS\rangle = \alpha|d^1\rangle + \beta|d^2\underline{L}\rangle. \tag{6}$$

For CTI/NCTEI, because $|d^1\rangle$ has lower (higher) energy than $|d^2\underline{L}\rangle$, α should be larger (smaller) than β and they coincide with each other when $\Delta_{CT} = 0$. There results are summarized in Fig. 2(e). The fact that $|d^2\underline{L}\rangle$ has lower energy than $|d^1\rangle$ in NCTEI has huge impact on the magnetic exchange coupling. To see this point, we will use two different methods: a perturbative treatment of cluster models and a mean-field treatment of bipartite square lattice model.

B. Perturbative treatment of cluster models

Here we study the standard 90° and 180° $d-p-d$ geometry shown in Figs. 3(a) and 3(b), the parameters marked there are inherited from Eq. (1).

It is well known that Hubbard model preserves both U(1) charge and SU(2) spin symmetry, with electron filling given, all the states can be classified by the spin quantum number S. For Hubbard model with $S = 0$ and $S = 1$ sectors available, it is convenient to define J as the energy difference,

$$J = E_{S=1}^{(0)} - E_{S=0}^{(0)}, \tag{7}$$

where $E_{S=1}^{(0)}$, $E_{S=0}^{(0)}$ are the energy of ground state in the $S = 1$ and $S = 0$ sector. So $J < 0$ (>0) indicates a FM (AFM) exchange coupling.

Here we consider two extremes: $\Delta_{CT} \gg 0$ (but still in the CTI regime) and $\Delta_{CT} \ll 0$ (but avoiding fully empty p orbital) so a perturbative treatment is possible. Table I lists the main results and the derivation can be found in Appendix B. From Table I, it is clear that the exchange coupling mechanism has different behavior at these two extremes. For $\Delta_{CT} \gg 0$, J of both 180° and 90° cases are quartic power of t^{dp} and 180° gives strong AFM while 90° gives weak FM, reflecting the GKA semiempirical rules. For $\Delta_{CT} \ll 0$, J is no longer quartic, but quadratic power of t^{dp} in 180° cases and is not a function of t^{dp} but proportional to $j_H^{(p)}$ in 90° case. Thus 180° gives weak AFM while 90° gives strong FM in this situation, totally reversing the GKA semiempirical rules.

Such a paradigm shift stems from the fact that $|d^2\underline{L}\rangle$ has larger weight than $|d^1\rangle$ in NCTEI as illustrated by Eq. (6). For 180° with $\Delta_{CT} \gg 0$, the zeroth-order state for $S = 0$ and

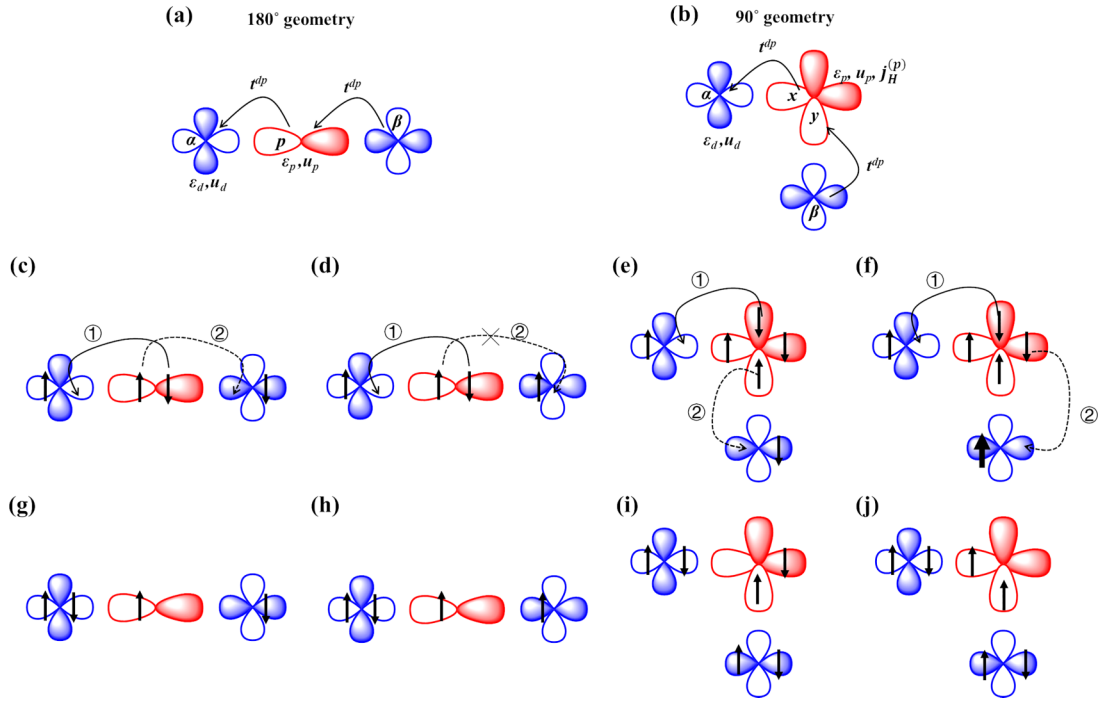

FIG. 3. (a) 180° $d-p-d$ geometry and (b) 90° geometry used in cluster model. Zeroth-order state of (c) $S=0$ sector for 180° with $\Delta_{\mathrm{CT}} \gg 0$ (d) $S=1$ sector for 180° with $\Delta_{\mathrm{CT}} \gg 0$ (e) $S=0$ sector for 90° with $\Delta_{\mathrm{CT}} \gg 0$ (f) $S=1$ sector for 90° with $\Delta_{\mathrm{CT}} \gg 0$ (g) $S=0$ sector for 180° with $-u_p < \Delta_{\mathrm{CT}} \ll 0$ (h) $S=1$ sector for 180° with $\Delta_{\mathrm{CT}} \ll 0$ (i) $S=0$ sector for 90° with $\Delta_{\mathrm{CT}} \ll 0$ (j) $S=1$ sector for 90° with $\Delta_{\mathrm{CT}} \ll 0$. The labels ① and ② represent the virtual electron hopping process between d and p orbital.

$S=1$ have the same energy, as shown in Figs. 3(c) and 3(d). Since hopping process marked by ① are allowed for both $S=0$ and $S=1$, they have the same energy at this level of perturbation [see Figs. 3(g) and 3(h)]. The energy difference between $S=0$ and $S=1$ comes from the second-order perturbation where hopping process ② is allowed in $S=0$ while forbidden in $S=1$ [see Figs. 3(c) and 3(d)]. In this way, $S=0$ has lower energy than $S=1$ with J positive and about $(t^{dp})^4$. The situation is almost the same for 90° with $\Delta_{\mathrm{CT}} \gg 0$, but when ① and ② [see Figs. 3(e) and 3(f)] are finished, the state in Fig. 3(i) has higher energy than Fig. 3(j) by the amount of $j_H^{(p)}$, which gives a weak negative J.

If $\Delta_{\mathrm{CT}} \ll 0$, the zeroth-order state for both geometry will be changed. For 180°, it is Figs. 3(g) and 3(h) for $S=0$ and $S=1$ sector, which is just the state with ① finished in Figs. 3(c) and 3(d) (here we insist $\Delta_{\mathrm{CT}} > -u_p$ to void empty p orbital in the zeroth-order state). Again, at the zeroth order, $S=0$ and $S=1$ have the same energy, but at the first order we can see the energy difference as hopping process ② is allowed in Fig. 3(g) but forbidden in Fig. 3(h). That is the reason why J is $\sim(t^{dp})^2$. While for 90° case, there is already an energy difference at the zeroth order as seen from Figs. 3(i) and 3(j), which gives a negative $J \sim j_H^{(p)}$ as listed in Table I.

From the above discussion, it is clear that negative charge-transfer energy states lead to strong FM in the system. To further see the effect of band structure introduced by t^d, t^p, next we are going to study the phase diagram on two-dimensional (2D) bipartite square lattice at the mean-field level [49].

C. Mean-field study of bipartite square lattice

The bipartite square lattice is shown in Fig. 4(a). To study both FM and (π, π)-AFM state, a $\sqrt{2} \times \sqrt{2}$ supercell was applied. Here we fixed $\varepsilon_d = 0$ and varied ε_p. For hopping terms, t^d and t^d were fixed to be –0.5 eV (i.e., $W_d = W_p =$ 4 eV) and t^{dp} was varied. For the interacting parameters, we fixed $u_d = 9$ eV and $u_p = 6$ eV so $u_d/|t^d|$ and $u_p/|t^p|$ is 18 and 12. With each d/p orbital contributes 1/2 electrons, the total electron filling (N_e) in the supercell is 6.0. A 100 × 100 k-point grid was applied for integration and the temperature was set to 30 K. During the mean-field simulation, 10 initial

TABLE I. Results of J for 180° and 90° $d-p-d$ geometry for CTI and NCTI at extreme condition, here Δ_{CT} is defined by Eq. (5).

Geometry	$\Delta_{\mathrm{CT}} \gg 0$	$\Delta_{\mathrm{CT}} \ll 0$
180°	$\frac{4(t^{dp})^4}{\Delta_{\mathrm{CT}}^2}\left(\frac{1}{u_d}+\frac{2}{2\Delta_{\mathrm{CT}}+u_p}\right)$ $(\Delta_{\mathrm{CT}} < u_d)$	$\frac{4(t^{dp})^2}{\Delta_{\mathrm{CT}}+u_p}$ $(\Delta_{\mathrm{CT}} > -u_p)$
90°	$-\frac{4(t^{dp})^4}{(\Delta_{\mathrm{CT}}-2u_p+5j_H^{(p)})^2}\left(\frac{1}{(2\Delta_{\mathrm{CT}}-3u_p+8j_H^{(p)})-j_H^{(p)}}-\frac{1}{(2\Delta_{\mathrm{CT}}-3u_p+8j_H^{(p)})+j_H^{(p)}}\right)$	$-2j_H^{(p)}$

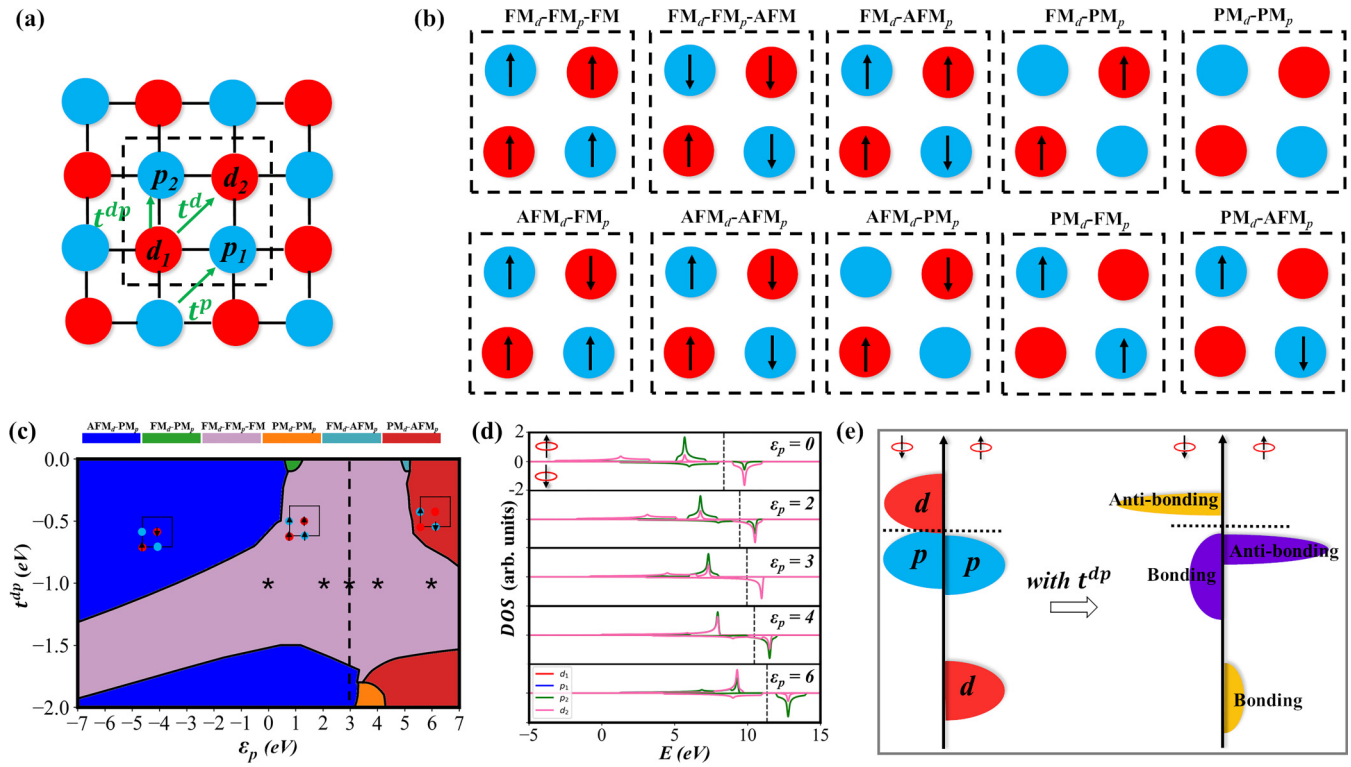

FIG. 4. (a) Bipartite square lattice. The blue/red filled circle represents p/d orbital. The dashed square represents a $\sqrt{2} \times \sqrt{2}$ supercell. (b) Initial magnetic phases used in the mean-field calculation. (c) Phase diagram, here $\varepsilon_d = 0$, $u_d = 9$ eV, and $u_p = 6$ eV. (d) Partial DOS of points marked by stars in (c). (e) Schematic representation of hybridization gap.

magnetic phases were applied as shown in Fig. 4(b). The label "FM$_d$-FM$_p$-FM" means that both d and p sublattices form FM order and the whole magnetic order is FM. The threshold for charge self-consistency was 10^{-5}.

The calculated phase diagram is shown in Fig. 4(c). When ε_p is far below ε_d, it is d orbital that is active, which forms AFM order (as d is half filled). Therefore, the ground state is AFM$_d$-PM$_p$. On the contrary, when ε_p is much higher than ε_d, it is p orbital that is active, which gives a PM$_d$-AFM$_p$ ground state (as p is also treated as correlated orbital here). These two extreme cases are in accordance with the perturbative treatment of cluster model. What is intriguing is that FM$_d$-FM$_p$-FM straddles with intermediate ε_p and such a phase is roughly symmetric with respect to $\varepsilon_p = 3$ eV (or $\Delta_{\mathrm{CT}} = 0$ eV) as shown in Fig. 4(c). Such a phenomena can be understood as follows: with Δ_{CT} close to zero, either d or p orbital is away from fully occupied (especially all long-range magnetic orders are metallic when $\Delta_{\mathrm{CT}} = 0$ eV), which gives the possibility of FM order in both d and p orbitals. This is further confirmed by the fact that with large $-t^{dp}$, FM$_d$-FM$_p$-FM can sustain in region with large $|\varepsilon_p|$. Although FM$_d$-FM$_p$-FM at $\varepsilon_p = 3$ eV is metallic, insulating phase can be find away from $\varepsilon_p = 3$ eV. Figure 4(d) displays the spin resolved partial density of state (DOS) of FM$_d$-FM$_p$-FM phase for different ε_p [marked as stars in Fig. 4(c) where t^{dp} was fixed to –1.0 eV]. When ε_p is close to 3 eV, the system remains metallic. Nevertheless, when ε_p is 0 or 6 eV, the system is insulating. The mechanism for gap opening at $\varepsilon_p = 0$ eV is orbital hybridization, as illustrated in Fig. 4(e). With $\varepsilon_p = 0$ eV, $\varepsilon_{d,\downarrow} - \varepsilon_{p,\downarrow} \sim (W_d + W_p)/2$, so a moderate t^{dp} is able to open a large gap. However, if $\varepsilon_{d,\downarrow} - \varepsilon_{p,\downarrow} \gg (W_d + W_p)/2$, then the system will become a CTI, which prefers AFM$_d$-PM$_p$ order. Meanwhile, if $\varepsilon_{d,\downarrow} - \varepsilon_{p,\downarrow} \ll (W_d + W_p)/2$, then huge t^{dp} is required to open a gap, which will drive the system to other magnetic orders.

In summary, with Δ_{CT} negative or close to zero, the exotic FM$_d$-FM$_p$-FM emerges and a moderate t^{dp} can open a hybridization gap under the condition that $\varepsilon_{d,\downarrow} - \varepsilon_{p,\downarrow} \sim (W_d + W_p)/2$.

D. Ferromagnetic negative charge transfer energy insulator

Here we would like to call the ferromagnetic insulator like Fig. 4(e) "ferromagnetic negative charge-transfer energy insulator". Compared with FM CTI or MHI, here the origin of gap is due to d-p orbital hybridization [50], rather than electron-electron correlations. Due to d-p hybridization, the obtaining bonding and antibonding orbitals in spin down channel of Fig. 4(e) are no longer localized d/p orbitals, but magnetic molecular orbitals (MMOs), which are linear combination of local d and p orbitals. These MMOs are nonlocal with large orbital spreadings and can reproduce the origin idea of Khomskii and Sawatzky [43]. Taken a bipartite square lattice formed by M and L [shown in Fig. 5(b)] as example, both the nearest (NN) and next-nearest neighbor (NNN) superexchange interaction (J_1 and J_2) are emergent FM according to Khomskii and Sawatzky [43]. In FNCTEI here, we are treating a MMO whose orbital spreading has a characteristic length covering both NN and NNN. Within this characteristic length, all exchange coupling should be FM and thus both J_1

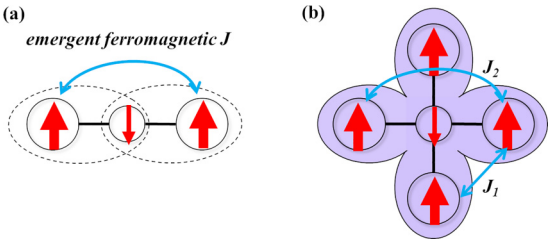

FIG. 5. (a) Schematic representation of emergent FM exchange coupling mediated by hole on L p orbital. The black dashed line represent strong AFM direct exchange coupling between M and L. (b) Magnetic molecular orbital with wave function extending over next-nearest neighbors, both J_1 and J_2 are FM. The large (small) circle represent M d (L p) orbital.

and J_2 are FM, in accordance with argument from Khomskii and Sawatzky [43].

To find FNCTEI in real materials, a multiband extension is required to take multiple d filling into consideration. At first glance, it is reasonable to assume that all d (p) orbitals have the same ε_d (ε_p) as the orbital splitting induced by point charges are generally small, therefore $j_H^{(d)}$ ($j_H^{(p)}$) will align spin polarization on d (p) shell to high spin configuration, making the picture in Fig. 4(e) still valid except that Fermi level can be higher. This makes two major differences. Firstly, from Fig. 4(e), it is clear that not all d filling can give a gap order so we have the idea of "ideal d filling": only at this d filling that the system exhibits a gap. When the d filling departs from "ideal d filling", extra electron/hole then enters antibonding/bonding orbitals of spin down channel, which destabilize the FM order. In this sense, the FM and gap order are synergistic in FNCTEI. Secondly, we need to consider another case when parts of d orbitals are empty. A simple calculation show that when Δ_{CT} is close to zero or negative, FM_d-FM_p-AFM will be the preferred ground state rather than FM_d-FM_p-FM (see Appendix C). It is further noted that if the local spin of M is 1/2, the coupling between d and p can lead to bounded Zhang-Rice singlets [51], which makes the system nonmagnetic (NM). Therefore, it is better for local spin of M to be in the classical limit. In the following, we will consider this situation only.

Strictly speaking, a semiconductor with FM_d-FM_p-AFM order should be coined as "ferrimagnetism" rather than ferromagnetism. But such a ferrimagnetism is inborn in the mechanism of FNCTEI, which is slightly different from normal ferrimagnetism where the underlying local spins come from different transition metal cations. For this reason, we still use "ferromagnetism" in FNCTEI.

III. MATERIAL REALIZATION

A. Material candidate

In this section, we look for material realization of FNCTEI. In principle, FNCTEI can be realized in any lattice structure. Here we choose PbO-type ML monolayer as an example. Such a binary lattice structure resembles the bipartite square lattice: the M atoms form a square lattice with L positioned in the middle of each square, alternatively above or below the paper plane. What is more, it is shared by $ThCr_2Si_2$- [52–55] and ZrCuSiAs-family materials [56–59] and has been widely studied due to the raising of Fe-based superconductors [60–62].

For the M-L combinations, M = V, Cr, Mn and L = P, As, Sb are selected. In transition metal oxides, it is well known that Δ_{CT} systematically decreases with increasing atomic number or increasing formal valence of the metallic ions [63]. For 3d metal cations with high-oxidation state (Cr^{4+}, Co^{3+}, Ni^{3+}, Cu^{3+}), Δ_{CT} can be very small or even negative. Therefore, we choose L from pnictogen family, which gives M a high oxidation M^{3+}. Nitrogen is ignored on purpose for its strong ionicity and weak covalent bonding with M. As for the metal ions, V, Cr, Mn are chosen to tune d filling (n) $2 \sim 4$. Here we focus on 3d transition metals for two reasons: firstly, they are lighter than their 4d and 5d cousins, the relativistic effect will be much weaker and would not drive the FM phase to other phases like quantum spin liquid [64]. Secondly, 3d orbitals has large $j_H^{(d)}$ than 4d and 5d, therefore high spin state is favored when multiple spin configurations are possible [65].

B. Paramagnetic phase

We first turn to the PM phase. The fitted single-particle parameter is shown in Fig. 6 (see Appendix D for model parameters fitting). Figure 6(a) shows the band structure without t^{dp}. Due to the high oxidation state of M, p orbitals are entangled with d orbitals, making the ε_{dp} close to 0 or even negative. When L goes from P to Sb, both W_p and W_d get much wider (from 4 to 8 eV). Figure 6(b) displays the strength of different $p-d$ hybridization channels [see Fig. 11(c) and Table IV for the definition]. The $p-d-5$ is the strongest in the buckling-free case such as in CuO_2 plane, but now it becomes almost the smallest (only larger than $p-d-8$). Thus such buckling cannot be regarded as small geometrical deviation here. The strongest $p-d$ hybridization now becomes the $p-d-1$ channel, which can be as large as 1.1 eV in VP and CrP, such a large σ-type hopping is quite astonishing. The second largest comes from the $p-d-2$, which is 1.0 eV in VP and MnP. Such large multiple $p-d$ hybridizations are essential to open a hybridization gap considering the large W_p and W_d. Seen from Fig. 6(b), as L goes from P to Sb, nearly all $p-d$ hybridizations become smaller. Considering the fact that W_p and W_d get wider from P to Sb, it will be harder for MSb to open a gap than MP and MAs in the FM phase.

Table II lists all the interaction parameters for the nine ML monolayers. The u_d is the leading energy scale, which is around 9 eV. The u_p term, the Hubbard interaction on p shell, is smaller but can be as large as 6 eV. Therefore the electron-electron correlations on p are large enough to form a submagnetic order. With L becoming heavier, u_d and u_p undergo large decreases. For example, when we go from MnAs to MnSb, the reduction of u_d (u_p) is up to 2.8 eV (2.3 eV). On the contrary, the Hund's coupling $j_H^{(d)}$ and $j_H^{(p)}$ are less influenced by the elemental differences and remain values proximate to 0.60 and 0.40 eV. Another noticeable fact is the large intershell Coulomb interaction u_{dp}, such a value is much larger than that in iron-based superconductors [66] and comparable with that expected in cuprates [47].

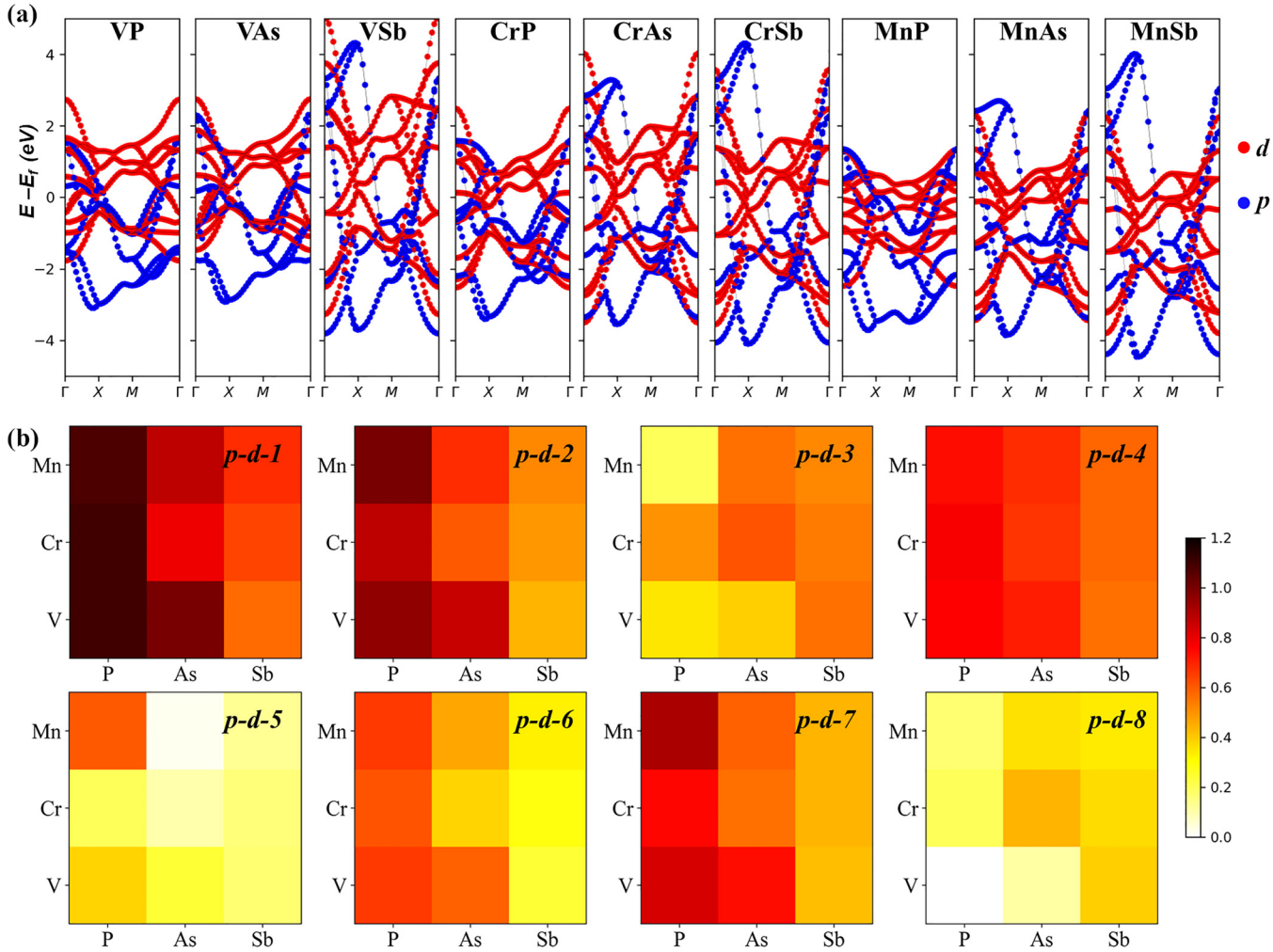

FIG. 6. (a) Band structure of PM M*L* monolayer without t^{dp}. The high symmetric k-path is Γ-X-M-Γ: (0.0, 0.0, 0.0)–(0.5, 0.0, 0.0)–(0.5, 0.5, 0.0)–(0.0, 0.0, 0.0). (b) Strength of t^{dp} in different $p-d$ hopping channel according to Table IV. To avoid phase dependence of t^{dp}, the phase is chosen so that all t^{dp} is positive. The unit is eV.

C. Ferromagnetic phase

1. n = 3

We first consider n = 3 and take CrAs as an example. Without $p-d$ hybridization, both spins are metallic as shown in Figs. 7(a) and 7(b). Compared with PM, ε_{dp} experiences huge changes: spin up p still highly entangle with d orbitals with several p bands higher than d bands while spin down p are well separated from d bands. With $p-d$ hybridization turning on, an indirect (direct) gap ∼0.35 eV (2.82 eV) opens in the spin up (down) channel as shown in Figs. 7(c) and 7(d), which makes FM CrAs monolayer a semiconductor. The origin of gap in different spin channels are different, as can be seen in the orbital projected band structure in Figs. 7(c) and 7(d). For spin up channel, both p and d orbitals have competing weights on the antibonding orbitals (mainly formed by p_x, p_y, d_{xz}, d_{yz}, d_{xy} as marked by light-yellow region) and several bonding orbitals, therefore the origin of gap here is orbital hybridization. Due to the large t^{dp} between p_x (p_y) and d_{xz} (d_{yz}) in channel $p-d-1$ and between p_x+p_y and d_{xy} in channel $p-d-2$, we can see a global gap in the spin up channel here. As for spin down channel, the gap is mainly assigned to the large ε_{dp}, as the antibonding (bonding) orbitals are dominated by the Cr 3d (As 3p) orbitals.

To see the magnetic molecular orbitals in the spin up channel, here we downfold the four antibonding magnetic Wannier functions (MWFs) and the results are plotted in Fig. 7(e). We find that the these MWFs consist of small CrAs clusters, rather than individual Cr or As atoms. To be specific, WF-1 is formed

TABLE II. Interaction parameters obtained from cRPA calculation. The unit is eV.

System	u_d	u'_d	$j_H^{(d)}$	u_p	u'_p	$j_H^{(p)}$	u_{dp}
VP	9.34	8.24	0.55	6.60	5.70	0.40	3.90
VAs	9.31	8.19	0.56	6.04	5.21	0.38	3.65
VSb	7.77	6.72	0.53	5.22	4.54	0.40	2.97
CrP	9.60	8.39	0.61	6.80	5.86	0.45	3.95
CrAs	8.89	7.67	0.61	6.08	5.25	0.46	3.37
CrSb	7.16	5.90	0.64	4.41	3.64	0.41	1.88
MnP	8.80	7.55	0.63	6.38	5.40	0.42	3.74
MnAs	9.39	8.07	0.66	6.34	5.47	0.47	3.66
MnSb	6.52	5.18	0.68	4.04	3.28	0.41	1.62

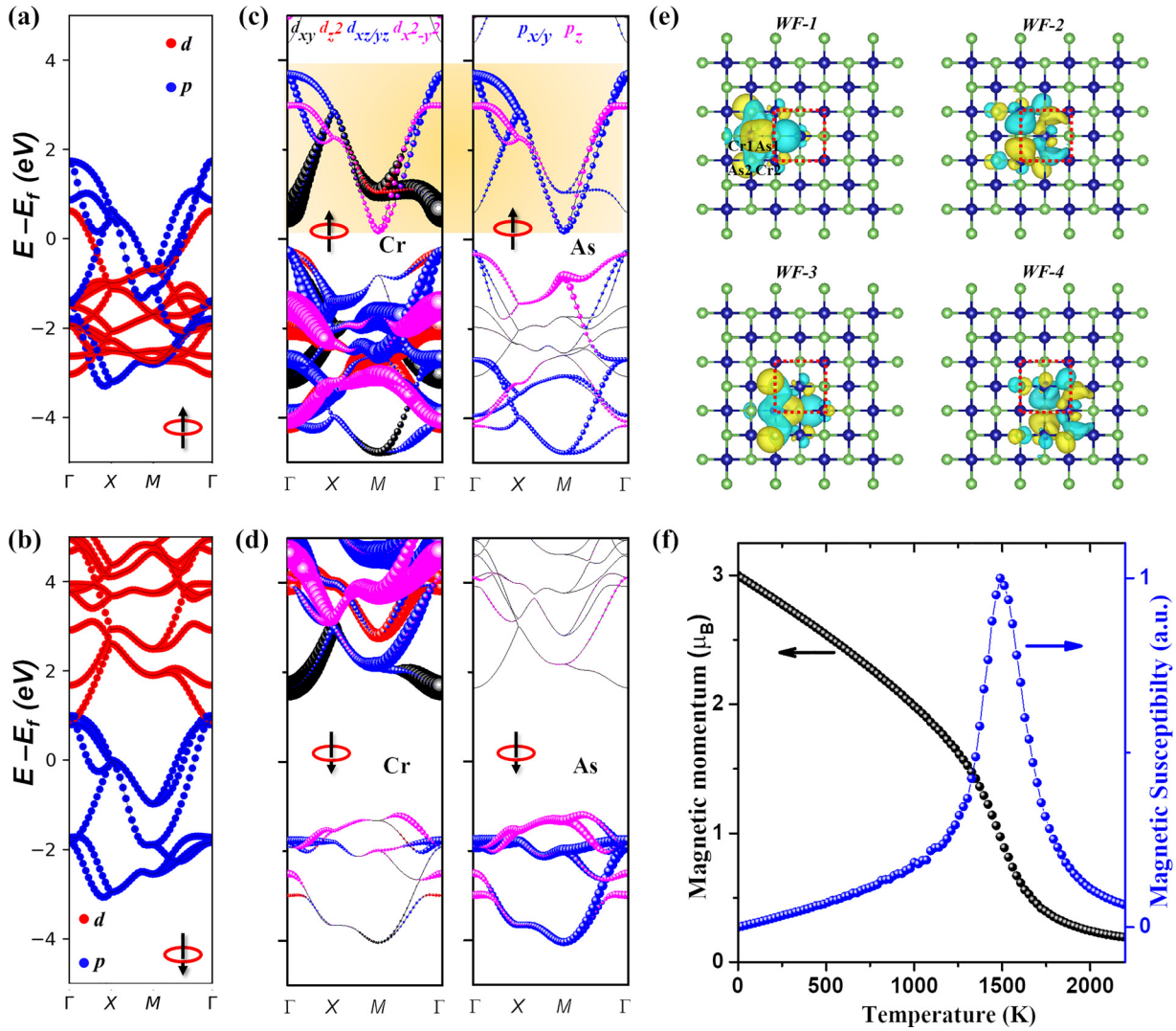

FIG. 7. [(a),(b)] Band structure of FM CrAs monolayer without t^{dp} for spin up/down channel. [(c),(d)] Orbital-resolved band structure for spin up/down channel. The left/right panel is for Cr/As respectively. The antibonding bands are marked by shaded yellow region in (c). In (c) and (d), different colors are used to stand for different atomic orbitals. (e) The four maximally localized Wannier functions for the antibonding orbitals for up spin. Red-dashed rectangle represents unit cell. (f) Evolution of effective magnetic moment (black) and magnetic susceptibility (blue) with respect to temperature.

by As1 p_x, Cr2 d_{xy}, and Cr1 d_{xz} while WF-2 is constituted by As1 p_y, Cr2 d_{yz}, and Cr1 d_{xy}, the other two WFs are equivalent to WF-1 and WF-2. Such a linear combination of orbitals is in accordance with the orbital composition of these bands depicted in Fig. 7(c). Two consequences come along: at first, p orbitals are populated with non-negligible holes, which gives a local magnetic moment –0.42 μ_B. Secondly, WFs are highly non-local and even cover the NNN Cr dimmers as plotted in Fig. 5(b). As discussed in Sec. II, both J_1 and J_2 should be FM.

According to energy mapping method (see Appendix E), the calculated J_1 and J_2 are –56.8 and –39.7 meV, both are ferromagnetic. In FM MHI and CTI, one expects the J to be rather short-range, the real surprise here is the magnitude of J_2 over J_1. The ratio J_2:$J_1 \sim 1/\sqrt{2}$:1 strongly implies linear-law scaling here. As both J_1 and J_2 are ferromagnetic, the magnetic ground state in such a square lattice is FM, in accordance with first-principles calculations. To determine T_c, classical Monte Carlo (MC) simulations are performed for a $32 \times 32 \times 1$ supercell based on Heisenberg Hamiltonian with J_1 and J_2 [67]. During FM-PM phase transition, magnetic susceptibility is calculated after the system reaches equilibrium at a given temperature, then T_c corresponds to the position of peak in magnetic susceptibility plot as shown in Fig. 7(f). As anticipated, an ultrahigh $T_c \sim 1500$ K is obtained. For 2D magnets, strong magnetic anisotropy energy (MAE) is needed to break the Hohenberg-Merin-Wagner theorem [68,69]. Since the highest occupied and lowest unoccupied orbital are mainly contributed by Cr $d_{xz/yz}$ and d_{xy} in spin up [see Fig. 7(c)], the orbital angular momentum difference ($||\Delta L_z||$) between these orbitals is 1, the prefer spin orientations of magnetic ions should be perpendicular to out-of-plane direction [70,71]. To verify this point, spin-orbit coupling (SOC) is taken into account in calculating the relative energies along (001) and (100) direction. It is (100) that is the easy axis, in accordance with orbital composition

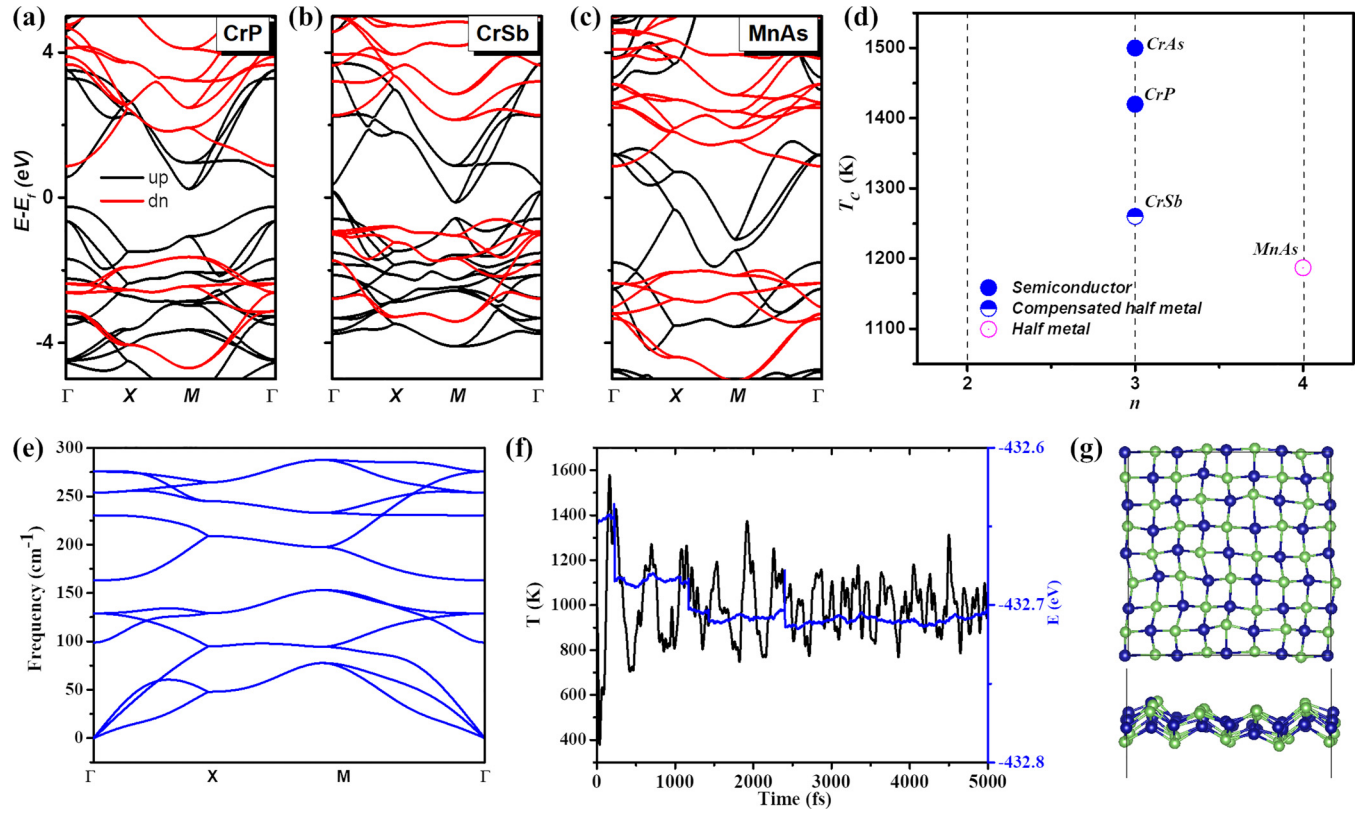

FIG. 8. [(a)–(c)] Band structure of FM CrP, CrAs, and MnAs monolayer. (d) A schematic phase diagram with respect to n. (e) Phonon spectrum of FM CrAs monolayer. (f) AIMD simulation of FM CrAs monolayer at 1000 K. (g) Top and side view of snapshot taken at 5 ps in AIMD simulation.

analysis. The calculated MAE is ∼0.37 meV/Cr, comparable to that in CrI_3 [72].

Now we count the electron filling. For the 18 electrons from Cr *d* and As *p* orbitals, 6 of them occupy the spin down bands and the left 12 electrons occupy the spin up bands. Seen from Fig. 7(c), these 12 electrons just fill bands up to the gap, which makes CrAs monolayer a semiconductor. In this sense, n = 3 is "ideal *d* filling" for such a lattice structure. Along this logic, CrP and CrSb monolayer should also belong to FNCTEI. Figures 8(a) and 8(b) show the band structure of FM CrP and CrSb monolayer. CrP monolayer resembles CrAs monolayer very much, which is also a small gap semiconductor with similar J_1 and J_2 as displayed in Table III. As for CrSb monolayer, because of larger W_p (W_d) and smaller t^{dp}, there is no global gap in Fig. 8(b), which makes CrSb monolayer a compensated half-metal. The calculated J_1 and J_2 (see Table III) of CrSb monolayer are smaller than CrAs and CrP monolayer, which gives a slightly lower $T_c \sim 1260$ K.

TABLE III. Exchange coupling strength and estimated T_c for different FM monolayers.

System	J_1 (meV)	J_2 (meV)	T_c (K)
CrP	−55.3	−36.2	1420
CrAs	−56.8	−39.7	1500
CrSb	−51.7	−25.8	1260
MnAs	−11.7	−32.5	1187

2. *Other fillings*

V*L* and Mn*L* can be regarded as one hole and electron doped Cr*L* per unit cell. Generally speaking, the extra charge doping will mediate FM exchange interactions between localized electrons, thus enhancing FM in the system. Nevertheless, from Fig. 4(e), the doping of holes (electrons) will decrease (increase) the occupation of bonding (antibonding) states, hence destabilizing the FM ground state (see Appendix G for details). The compete between these two factors may drive FM-AFM transition. To see the robustness of FM in FNCTEI, here we go on studying V*L* and Mn*L* monolayers. For V*L*, no magnetic orders are found, which maybe due to the small local magnetic moment on V. For MnL, FM solutions are found for MnAs while MnP and MnSb prefers AFM ground state [73]. The band structure of FM MnAs monolayer is depicted in Fig. 8(c), which is a half-metal. Table III lists J_1 and J_2 value of MnAs, it is clear that J_1 is much smaller than that in CrL, suggesting the destabilization effect is much stronger than itinerant enhancement. Such a phenomena can be taken as an indicator for FNCTEI in experiment. With large J_2 and larger magnetic moments, the T_c of MnAs monolayer is still over 1100 K.

Figure 8(d) summarizes the results for all the FM monolayers. For the "ideal *d* filling" n = 3, FNCTEI is found in CrAs and CrP monolayer, a nearby phase of FNCTEI is FM compensated half-metal, as found in CrSb monolayer. Deviating from ideal filling, no FM is obtained in n = 2 while FM half-metal phase is possible for n = 4, like in MnAs monolayer.

3. Stability of CrAs monolayer

The existence of bulk CrAs in nature [74,75] suggests the 1:1 stoichiometric ratio in CrAs monolayer is charge feasible. The stability of CrAs monolayer is confirmed by both phonon spectrum and *ab-initio* molecular dynamics (AIMD) [Figs. 8(e)–8(g)], in accordance with previous report [76]. The stability indicates an experimental preparation of CrAs monolayer is possible. With layered $BaCr_2As_2$ [54] and LaCrAsO [59] synthesized in experiments, CrAs monolayer can be obtained by either etching [77] or electrochemical reactions [78] from these layered materials. Another synthetic method is molecular-beam epitaxy and $BaZrO_3$ (001) [or MgO (001)] is a perfect substrate with 1:1 lattice match. Such a strategy has shown success in CoSb monolayer preparation [79,80].

IV. DISCUSSION AND CONCLUSIONS

Up to here, we illustrate the concept of FNCTEI and suggest the material realization in CrAs monolayer. FNCTEI are different from FM CTI and MHI from the following aspects: (1) *Origin of gap*. In FM CTI and MHI, the gap is originated from electron-electron correlation and is always large. However in FNCTEI, the gap comes from d-p hybridization, which is commonly small. Since strong orbital hybridization reflects the covalency nature of the system, FNCTEI thus lies at a special point where kinetic and interaction energy make peace. In this regard, HSE06 functional is a good functional to study FNCTEI from perspective of first-principles calculation (see Appendix A for a detailed discussion). (2) *Scaling behavior of J*. In FM CTI and MHI, the building motifs of magnetic interaction are local atomic orbitals with nearly fully filled p orbitals, so J decays quickly with respect to distance between magnetic pairs. However, in FNCTEI, the building motifs are MMOs, theirs large orbital extension gives slowly decaying J within MMOs and sudden drop beyond MMOs (see Appendix F). (3) *Response to electron/hole doping*. In FM CTI and MHI, electron or hole doping tends to bring itinerant FM into the system and thus enhancing FM. While in FNCTEI, extra electron or hole will destabilize the pristine FM phase (see Appendix G for details). This may explain why LaCrAsO occupies AFM order [59]. Replacing La by Sr, part of Cr^{2+} will become Cr^{3+} and a AFM-FM transition may be observed.

It is well known that W_d and W_p will becomes larger as the system grows from 2D to 3D, thus FNCTEI is harder to be found in pure 3D than 2D materials. In 2D materials, because the band width and d-p hybridization are easily tuned by external strain, a FM metal-to-insulator phase transition can be achieved via strain engineering. Such a phenomena would be hard to observe in FM CTI or FM MHI. What is more, as p can be higher than d orbitals in FNCTEI, in view of band topology, this gives band inversion and nontrivial topology [81]. Therefore, FNCTEI is also a good platform to study quantum anomalous Hall effect.

We believe our paper also adds a member to the family of materials with negative charge-transfer energy. Nonmagnetic negative charge-transfer energy insulator has been reported in $NaCuO_2$ [41], both nonmagnetic negative charge-transfer energy insulator and metal in $RNiO_3$ (R is a rare earth element) [42], and ferromagnetic negative-charge transfer energy metal in CrO_2 [39], $SrCoO_3$ [40]. Here CrAs monolayer represents an ideal example of ferromagnetic negative charge-transfer energy insulator.

In conclusion, here a different type of FMSC is proposed: FNCTEI. It breaks the GKA semiempirical rules and the corresponding T_c can be much higher than room temperature. Through first-principles calculation with hybrid functional, CrAs monolayer is mapped out to be a typical FNCTEI, which has a band gap around 0.35 eV and a high T_c about 1500 K. Due to its exotic physical properties, it is envisioned that FNCTEI will arouse broad interest in condensed matter physics.

ACKNOWLEDGMENTS

Z.L. thanks A. van Roekeghem and H. Jiang for helpful discussion. This work is supported by NSFC (Grants No. 22288201, No. 12174356, and No. 22073086), National Key R&D Program of China (Grants No. 2017YFA0204904, No. 2016YFA0200604), Youth Innovation Promotion Association CAS (2019441) and the Start-up Funding from Westlake University. We thank Supercomputing Center at USTC for providing the computing resources.

APPENDIX

In the Appendix, we provide more details of the calculation and results to support the discussion in the main text. In Appendix A, we make a brief introduction to the calculation details, especially we aim at explaining why the hybrid functional is a good choice for the study of FNCTEI. In Appendix B, we derive the results in Table I. This section includes three subsections. In Appendix C, we extend the mean-field results to situations with empty d orbitals. In Appendix D, we talk about the model parameter calculation, including the single-particle part and the interacting part. In Appendix E, we discuss the energy mapping method where the exchange coupling strength J are calculated. In Appendix F, we focus on the scaling behavior of J, here a large supercell is applied and exchange coupling strengths up to J_4 are obtained. In Appendix G, we study the response of J to both hole and electron doping.

APPENDIX A: CALCULATION DETAILS AND FUNCTIONAL DEPENDENCE

Our first-principles calculations were performed on density functional theory implemented in the Vienna *Ab Initio* Simulation Package (VASP) [82]. For geometric optimization and electronic property calculations, a plane-wave cutoff 600 eV was used. The energy convergence criterion was 10^{-6} eV and the residual force was 0.01 eV/Å. The Brillouin zone integration was carried out with $12 \times 12 \times 1$ k-point sampling for paramagnetic phase with PBE functional [83]. The PBE functional was also applied in phonon spectrum calculation [84] and AIMD simulation at 1000 K [85].

To study the magnetic phase, a few remarks should be made here. As discussed in the main text both p and d orbitals are close to Fermi level, a suitable theory describing FNCTEI

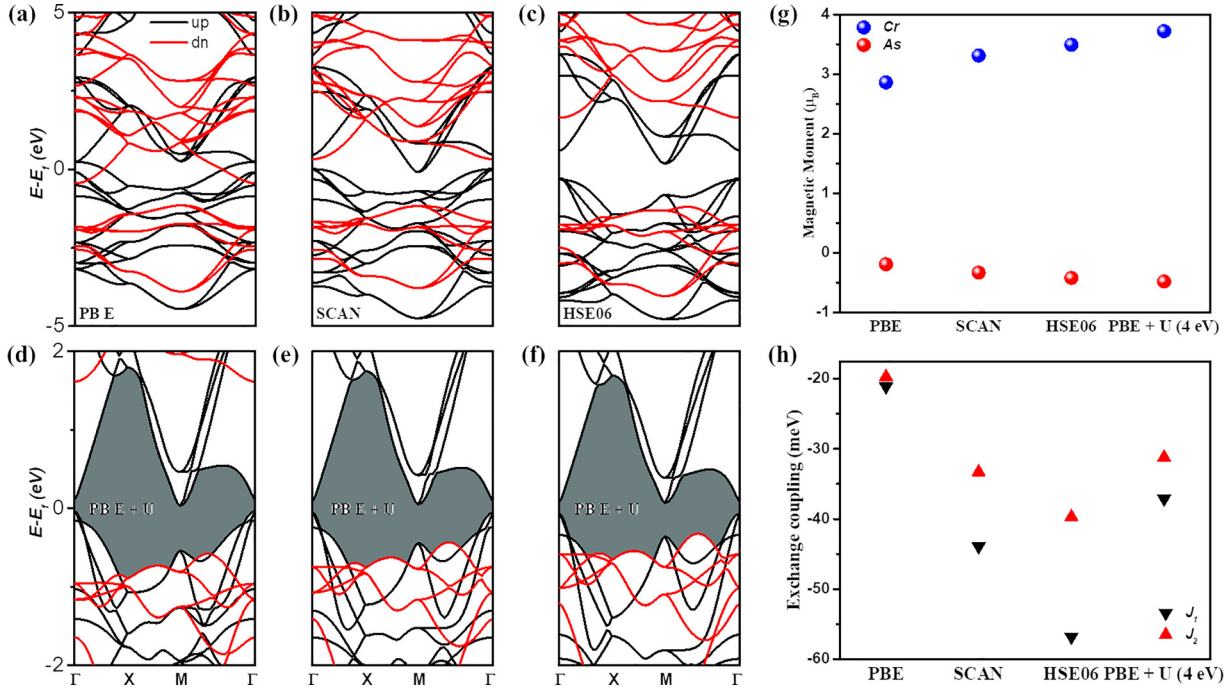

FIG. 9. Calculated band structure of FM CrAs monolayer by (a) PBE functional, (b) SCAN functional, (c) HSE06 functional, and PBE + U method with (d) U = 3 eV, (e) U = 4 eV, and (f) U = 5 eV. (g) Dependence of magnetic moments of Cr and As with respect to different functionals and methods. (h) Dependence of exchange coupling J_1 and J_2 with respect to different functionals and methods.

should contain both interaction terms of both p and d orbitals. Especially, intershell interaction term is also needed to describe the charge correlations between d and p shells. Such a feature impose strong constraint on theoretical approach. In this sense, Heyd-Scuseria-Ernzerhof (HSE) functional [86,87] should be better than functional such as PBE, SCAN [88,89] or embedding method such as standard PBE + U [90,91].

To see the performance of different functionals and methods, we used them to study the CrAs FM phase. The results are summarized in Fig. 9. We first consider the sequence of PBE, SCAN and HSE06 as they lie at the second, third, and fourth rungs in the Jacob's ladder [92]. From Fig. 9(a), PBE gives a itinerant FM phase, so the local magnetic moments on Cr and As are the smallest among these four functionals and methods. The SCAN functional is able to create a local gap for spin up channel, but is unable to create global gap [Fig. 9(b), which makes CrAs a compensated FM metal]. Due to the existence of local gap, the local magnetic moments on Cr and As are much larger than that of PBE. Finally, HSE06 corrects the gap to a global one ~ 0.35 eV of spin up as shown in Fig. 9(c). As the functional becomes more and more advanced as we go from PBE to SCAN to HSE06, the true magnetic properties are gradually approached.

Next we consider the PBE + U method, here we only added Hubbard U term on Cr 3d shell and U = 3, 4, and 5 eV were applied to see the trend. Different U give a similar band structure as shown in Figs. 9(d)–9(f). As in HSE06 functional, a global gap is opened, but the band gap of spin up is only about 0.10 eV, much smaller than HSE06. The local magnetic moments of PBE + U method is much larger than SCAN and HSE06, indicating PBE + U tends to localize the electrons. At the same time, the insufficient description of kinetic energy of PBE + U gives a gap much smaller than than HSE06.

Combining the above two considerations together, it can be seen that HSE06 functional is the best functional (method) in describing both the valency and correlation character of FNCTEI. In the main text, the magnetic properties, including band structure, exchange coupling strength estimation and magnetic isotropic energy are calculated using the HSE06 functional. For the band calculation, the wave function is preconverged from PBE functional with a $9 \times 9 \times 1$ k-point mesh for FM phase, and the eigenstates are then solved along three high symmetric k-paths Γ (0.0, 0.0, 0.0)-X (0.5, 0.0, 0.0)-M (0.5, 0.5, 0.0)-Γ (0.0, 0.0, 0.0) with 20 points along each path. For the $2 \times 2 \times 1$ supercell used in exchange coupling strength estimation, we use a $12 \times 12 \times 1$ k-point sampling.

APPENDIX B: DERIVATION OF TABLE I

In this Appendix, we derive the results in Table I. In the following, the many-body state is written in the occupation representation. For 180° geometry shown in Fig. 3(a), the orbital order in the many-body state is d_α, p, and d_β. For 90° geometry shown in Fig. 3(b), the orbital order in the many-body state is d_α, p_x, p_y, and d_β. For each orbital, $|0,0\rangle$, $|\uparrow,0\rangle$, $|0,\downarrow\rangle$, $|\uparrow,\downarrow\rangle$ represent empty state, singly occupied state with spin up/down and doubly occupied state, respectively.

For a 2×2 matrix as follows:

$$\begin{bmatrix} H_{00} & T_{01} \\ T_{10} & H_{11} \end{bmatrix}. \tag{B1}$$

If we are interested in H_{00} and the eigenvalues of H_{11} are separated from H_{00} with a large gap, we can integrate degrees

of freedom of H_{11} out by

$$H_{\text{eff}} = H_{00} + T_{01}(\varepsilon - H_{11})^{-1}T_{10}, \tag{B2}$$

where ε is the eigenvalue of original 2×2 matrix. Approximating ε by eigenvalue of H_{00}, we obtain the effective matrix describing the degrees of freedom of H_{00}.

If there are one more high-energy scale characterized by H_{22}, in other words we are treating a 2×2 matrix as follows:

$$\begin{bmatrix} H_{00} & T_{01} & H_{02} \\ T_{10} & H_{11} & H_{12} \\ T_{20} & H_{21} & H_{22} \end{bmatrix}. \tag{B3}$$

By recursively applying Eq. (B2), we have

$$H_{\text{eff}} = H_{00} + T_{01}[\varepsilon - (H_{11} + T_{12}(\varepsilon - H_{22})^{-1}T_{21})]^{-1}T_{10}. \tag{B4}$$

The following approximation is often used to simplify the calculation of inverse of a matrix:

$$(A - B)^{-1} \sim A^{-1} + A^{-1}BA^{-1}, \tag{B5}$$

where the eigenvalues of A separate from that of B by a large gap.

1. 180° geometry

a. S = 0 sector

There are 6 states in the $S = 0$ sector,

$$\Big\{ \frac{1}{\sqrt{2}}(|\uparrow,0;\uparrow,\downarrow;0,\downarrow\rangle - |0,\downarrow;\uparrow,\downarrow;\uparrow,0\rangle),$$
$$\frac{1}{\sqrt{2}}(|\uparrow,\downarrow;\uparrow,0;0,\downarrow\rangle - |\uparrow,\downarrow;0,\downarrow;\uparrow,0\rangle),$$
$$\frac{1}{\sqrt{2}}(|\uparrow,0;0,\downarrow;\uparrow,\downarrow\rangle - |0,\downarrow;\uparrow,0;\uparrow,\downarrow\rangle),$$
$$|\uparrow,\downarrow;\uparrow,\downarrow;0,0\rangle,\ |0,0;\uparrow,\downarrow;\uparrow,\downarrow\rangle,\ |\uparrow,\downarrow;0,0;\uparrow,\downarrow\rangle \Big\}.$$

For example, the many-body state $\frac{1}{\sqrt{2}}(|\uparrow,0;\uparrow,\downarrow;0,\downarrow\rangle - |0,\downarrow;\uparrow,\downarrow;\uparrow,0\rangle)$ describes a singlet between d_α and d_β and a doubly occupied p.

For $\Delta_{\text{CT}} \gg 0$, the energy of $\frac{1}{\sqrt{2}}(|\uparrow,0;\uparrow,\downarrow;0,\downarrow\rangle - |0,\downarrow;\uparrow,\downarrow;\uparrow,0\rangle)$ is $2\varepsilon_d + 2\varepsilon_p + u_p$, which we will take as reference energy. Accordingly, the energy of $\frac{1}{\sqrt{2}}(|\uparrow,\downarrow;\uparrow,0;0,\downarrow\rangle - |\uparrow,\downarrow;0,\downarrow;\uparrow,0\rangle)$ is $\varepsilon_d - \varepsilon_p + u_d - u_p$, which is just the charge transfer energy Δ_{CT} defined in Eq. (5). At the same, the state $|\uparrow,\downarrow;\uparrow,\downarrow;0,0\rangle$ has energy u_d. If $\Delta_{\text{CT}} > u_d$, the system lies at MHI regime and the lowest excitation energy is u_d rather than Δ_{CT}. Here we are interested in CTI regime and assume $\Delta_{\text{CT}} < u_d$.

The configuration interaction matrix for the above basis is

$$\begin{bmatrix} 0 & t^{dp} & t^{dp} & 0 & 0 & 0 \\ t^{dp} & \Delta_{\text{CT}} & 0 & -\sqrt{2}t^{dp} & 0 & -\sqrt{2}t^{dp} \\ t^{dp} & 0 & \Delta_{\text{CT}} & 0 & -\sqrt{2}t^{dp} & -\sqrt{2}t^{dp} \\ 0 & -\sqrt{2}t^{dp} & 0 & u_d & 0 & 0 \\ 0 & 0 & -\sqrt{2}t^{dp} & 0 & u_d & 0 \\ 0 & -\sqrt{2}t^{dp} & -\sqrt{2}t^{dp} & 0 & 0 & 2\Delta_{\text{CT}} + u_p \end{bmatrix}. \tag{B6}$$

Here we are interested in $\varepsilon \sim 0$ and take

$$H_{00} = 0,$$
$$H_{11} = \begin{bmatrix} \Delta_{\text{CT}} & 0 \\ 0 & \Delta_{\text{CT}} \end{bmatrix},$$
$$H_{22} = \begin{bmatrix} u_d & 0 & 0 \\ 0 & u_d & 0 \\ 0 & 0 & 2\Delta_{\text{CT}} + u_p \end{bmatrix}. \tag{B7}$$

By setting $\varepsilon = 0$, Eq. (B4) becomes

$$H_{\text{eff}} \approx H_{00} - T_{01}\left(H_{11} - T_{12}H_{22}^{-1}T_{21}\right)^{-1}T_{10}. \tag{B8}$$

By taking A = H_{11} and B = $T_{12}H_{22}^{-1}T_{21}$ in Eq. (B5), we can further simplify the above equation as

$$H_{\text{eff}} \approx H_{00} - T_{01}H_{11}^{-1}T_{10} - T_{01}H_{11}^{-1}T_{12}H_{22}^{-1}T_{21}H_{11}^{-1}T_{10}. \tag{B9}$$

By plugging all the matrix into above equation, we have

$$H_{\text{eff}} \approx -\frac{2t^{dp}}{\Delta_{\text{CT}}} - \frac{4(t^{dp})^2}{\Delta_{\text{CT}}}\left(\frac{(t^{dp})^2}{u_d} + \frac{2(t^{dp})^2}{2\Delta_{\text{CT}} + u_p}\right). \tag{B10}$$

For $\Delta_{\text{CT}} \ll 0$ (but we insist $\Delta_{\text{CT}} > -u_p$), state $\frac{1}{\sqrt{2}}(|\uparrow,\downarrow;\uparrow,0;0,\downarrow\rangle - |\uparrow,\downarrow;0,\downarrow;\uparrow,0\rangle)$ and $\frac{1}{\sqrt{2}}(|\uparrow,0;0,\downarrow;\uparrow,\downarrow\rangle - |0,\downarrow;\uparrow,0;\uparrow,\downarrow\rangle)$ will have lower energy than $\frac{1}{\sqrt{2}}(|\uparrow,0;\uparrow,\downarrow;0,\downarrow\rangle - |0,\downarrow;\uparrow,\downarrow;\uparrow,0\rangle)$, at this time, we arrange the

basis as

$$\left\{\frac{1}{\sqrt{2}}(|\uparrow,\downarrow;\uparrow,0;0,\downarrow\rangle-|\uparrow,\downarrow;0,\downarrow;\uparrow,0\rangle),\ \frac{1}{\sqrt{2}}(|\uparrow,0;0,\downarrow;\uparrow,\downarrow\rangle-|0,\downarrow;\uparrow,0;\uparrow,\downarrow\rangle),\ |\uparrow,\downarrow;0,0;\uparrow,\downarrow\rangle,\right.$$
$$\left.\frac{1}{\sqrt{2}}(|\uparrow,0;\uparrow,\downarrow;0,\downarrow\rangle-|0,\downarrow;\uparrow,\downarrow;\uparrow,0\rangle),\ |\uparrow,\downarrow;\uparrow,\downarrow;0,0\rangle,\ |0,0;\uparrow,\downarrow;\uparrow,\downarrow\rangle\right\}$$

and then the configuration interaction matrix becomes

$$\begin{bmatrix} \Delta_{\mathrm{CT}} & 0 & -\sqrt{2}t^{dp} & t^{dp} & -\sqrt{2}t^{dp} & 0 \\ 0 & \Delta_{\mathrm{CT}} & -\sqrt{2}t^{dp} & t^{dp} & 0 & -\sqrt{2}t^{dp} \\ -\sqrt{2}t^{dp} & -\sqrt{2}t^{dp} & 2\Delta_{\mathrm{CT}}+u_p & 0 & 0 & 0 \\ t^{dp} & t^{dp} & 0 & 0 & 0 & 0 \\ -\sqrt{2}t^{dp} & 0 & 0 & 0 & u_d & 0 \\ 0 & -\sqrt{2}t^{dp} & 0 & 0 & 0 & u_d \end{bmatrix}. \tag{B11}$$

Here we are interested in $\varepsilon \sim \Delta_{\mathrm{CT}}$ and take

$$H_{00}=\begin{bmatrix}\Delta_{\mathrm{CT}} & 0\\ 0 & \Delta_{\mathrm{CT}}\end{bmatrix},$$
$$H_{11}=\begin{bmatrix}2\Delta_{\mathrm{CT}}+u_p & 0\\ 0 & 0\end{bmatrix},$$
$$H_{22}=\begin{bmatrix}u_d & 0\\ 0 & u_d\end{bmatrix}. \tag{B12}$$

Plugging these terms into Eq. (B4), we have

$$H_{\mathrm{eff}}\approx\left[\Delta_{\mathrm{CT}}+\left(\frac{(t^{dp})^2}{\Delta_{\mathrm{CT}}}-\frac{2(t^{dp})^2}{\Delta_{\mathrm{CT}}+u_p}\right)\right]\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$$
$$+\left(\frac{(t^{dp})^2}{\Delta_{\mathrm{CT}}}-\frac{2(t^{dp})^2}{\Delta_{\mathrm{CT}}+u_p}\right)\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}. \tag{B13}$$

b. S = 1 sector

There are 9 states in the $S=1$ sector,

$$\left\{|0,\downarrow;\uparrow,\downarrow;0,\downarrow\rangle,\ |\uparrow,\downarrow;0,\downarrow;0,\downarrow\rangle,\ |0,\downarrow;0,\downarrow;\uparrow,\downarrow\rangle,\right.$$
$$\frac{1}{\sqrt{2}}(|\uparrow,0;\uparrow,\downarrow;0,\downarrow\rangle+|0,\downarrow;\uparrow,\downarrow;\uparrow,0\rangle),$$
$$\frac{1}{\sqrt{2}}(|\uparrow,\downarrow;\uparrow,0;0,\downarrow\rangle+|\uparrow,\downarrow;0,\downarrow;\uparrow,0\rangle),$$
$$\frac{1}{\sqrt{2}}(|\uparrow,0;0,\downarrow;\uparrow,\downarrow\rangle+|0,\downarrow;\uparrow,0;\uparrow,\downarrow\rangle),$$
$$\left.|\uparrow,0;\uparrow,\downarrow;\uparrow,0\rangle,\ |\uparrow,\downarrow;\uparrow,0;\uparrow,0\rangle,\ |\uparrow,0;\uparrow,0;\uparrow,\downarrow\rangle\right\}.$$

Notice that the first three states, second three states and the last three states are not connected via t^{dp}, therefore, we can work on a smaller subspace expanded by the last three states,

$$\{|\uparrow,0;\uparrow,\downarrow;\uparrow,0\rangle,\ |\uparrow,\downarrow;\uparrow,0;\uparrow,0\rangle,\ |\uparrow,0;\uparrow,0;\uparrow,\downarrow\rangle\}.$$

For $\Delta_{\mathrm{CT}}\gg 0$, the configuration interaction matrix in such a basis is

$$\begin{bmatrix}0 & t^{dp} & t^{dp}\\ t^{dp} & \Delta_{\mathrm{CT}} & 0\\ t^{dp} & 0 & \Delta_{\mathrm{CT}}\end{bmatrix}. \tag{B14}$$

Taking $\varepsilon=0$

$$H_{00}=0,$$
$$H_{11}=\begin{bmatrix}\Delta_{\mathrm{CT}} & 0\\ 0 & \Delta_{\mathrm{CT}}\end{bmatrix}, \tag{B15}$$

and making use of Eq. (B2), we have

$$H_{\mathrm{eff}}=-\frac{2(t^{dp})^2}{\Delta_{\mathrm{CT}}}. \tag{B16}$$

Combining Eqs. (B10) and (B16), J is calculated as

$$J=-\frac{2(t^{dp})^2}{\Delta_{\mathrm{CT}}}+\left[\frac{2(t^{dp})^2}{\Delta_{\mathrm{CT}}}+\frac{4(t^{dp})^2}{\Delta_{\mathrm{CT}}}\left(\frac{(t^{dp})^2}{u_d}+\frac{2(t^{dp})^2}{2\Delta_{\mathrm{CT}}+u_p}\right)\right]$$
$$=\frac{4(t^{dp})^2}{\Delta_{\mathrm{CT}}}\left(\frac{(t^{dp})^2}{u_d}+\frac{2(t^{dp})^2}{2\Delta_{\mathrm{CT}}+u_p}\right). \tag{B17}$$

For $\Delta_{\mathrm{CT}}\ll 0$, the basis is rearranged as

$$\{|\uparrow,\downarrow;\uparrow,0;\uparrow,0\rangle,\ |\uparrow,0;\uparrow,0;\uparrow,\downarrow\rangle,\ |\uparrow,0;\uparrow,\downarrow;\uparrow,0\rangle\},$$

and the configuration interaction matrix is

$$\begin{bmatrix}\Delta_{\mathrm{CT}} & 0 & t^{dp}\\ 0 & \Delta_{\mathrm{CT}} & t^{dp}\\ t^{dp} & t^{dp} & 0\end{bmatrix}. \tag{B18}$$

Taking $\varepsilon = \Delta_{\mathrm{CT}}$

$$H_{00} = \begin{bmatrix} \Delta_{\mathrm{CT}} & 0 \\ 0 & \Delta_{\mathrm{CT}} \end{bmatrix}, \quad H_{11} = 0, \tag{B19}$$

and making use of Eq. (B2), we have

$$H_{\mathrm{eff}} \approx \left(\Delta_{\mathrm{CT}} + \frac{(t^{dp})^2}{\Delta_{\mathrm{CT}}}\right)\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{(t^{dp})^2}{\Delta_{\mathrm{CT}}}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \tag{B20}$$

Since Eqs. (B13) and (B20) are both 2×2 matrices, we first need to diagonalize it and then J is calculated as

$$\begin{aligned} J &= \left(\Delta_{\mathrm{CT}} + \frac{2(t^{dp})^2}{\Delta_{\mathrm{CT}}}\right) \\ &\quad - \left[\Delta_{\mathrm{CT}} + 2\left(\frac{(t^{dp})^2}{\Delta_{\mathrm{CT}}} - \frac{2(t^{dp})^2}{\Delta_{\mathrm{CT}} + u_p}\right)\right] \\ &= \frac{4(t^{dp})^2}{\Delta_{\mathrm{CT}} + u_p}. \end{aligned} \tag{B21}$$

Since $\Delta_{\mathrm{CT}} + u_p$ is positive, Eq. (B21) still gives antiferromagnetic exchange coupling.

2. 90° geometry

a. S = 0 sector

There are 10 states in the S = 0 sector,

$$\left\{ \frac{1}{\sqrt{2}}(|\uparrow,0;\uparrow,\downarrow;\uparrow,\downarrow;0,\downarrow\rangle - |0,\downarrow;\uparrow,\downarrow;\uparrow,\downarrow;\uparrow,0\rangle), \right.$$
$$\frac{1}{\sqrt{2}}(|\uparrow,0;\uparrow,\downarrow;0,\downarrow;\uparrow,\downarrow\rangle - |0,\downarrow;\uparrow,\downarrow;\uparrow,0;\uparrow,\downarrow\rangle),$$
$$\frac{1}{\sqrt{2}}(|\uparrow,\downarrow;\uparrow,0;\uparrow,\downarrow;0,\downarrow\rangle - |\uparrow,\downarrow;0,\downarrow;\uparrow,\downarrow;\uparrow,0\rangle),$$
$$\frac{1}{\sqrt{2}}(|\uparrow,\downarrow;\uparrow,0;0,\downarrow;\uparrow,\downarrow\rangle - |\uparrow,\downarrow;0,\downarrow;\uparrow,0;\uparrow,\downarrow\rangle),$$
$$\frac{1}{\sqrt{2}}(|\uparrow,\downarrow;\uparrow,\downarrow;\uparrow,0;0,\downarrow\rangle - |\uparrow,\downarrow;\uparrow,\downarrow;0,\downarrow;\uparrow,0\rangle),$$
$$\frac{1}{\sqrt{2}}(|\uparrow,0;0,\downarrow;\uparrow,\downarrow;\uparrow,\downarrow\rangle - |0,\downarrow;\uparrow,0;\uparrow,\downarrow;\uparrow,\downarrow\rangle),$$
$$|\uparrow,\downarrow;\uparrow,\downarrow;\uparrow,\downarrow;0,0\rangle,\ |\uparrow,\downarrow;\uparrow,\downarrow;0,0;\uparrow,\downarrow\rangle,$$
$$\left. |\uparrow,\downarrow;0,0;\uparrow,\downarrow;\uparrow,\downarrow\rangle,\ |0,0;\uparrow,\downarrow;\uparrow,\downarrow;\uparrow,\downarrow\rangle \right\}.$$

Notice that the first four states are not connected to the last six states via t^{dp}, to simplify the discussion, we will work in the subspace spanned by the first four states.

For $\Delta_{\mathrm{CT}} \gg 0$, the configuration interaction matrix for the subspace is

$$\begin{bmatrix} 0 & t^{dp} & t^{dp} & 0 \\ t^{dp} & \Delta_{\mathrm{CT}} - 2u_p + 5j_H^{(p)} & 0 & t^{dp} \\ t^{dp} & 0 & \Delta_{\mathrm{CT}} - 2u_p + 5j_H^{(p)} & t^{dp} \\ 0 & t^{dp} & t^{dp} & 2\Delta_{\mathrm{CT}} - 3u_p + 9j_H^{(p)} \end{bmatrix}. \tag{B22}$$

Here we are interested in $\varepsilon \sim 0$ and put

$$H_{00} = 0, \quad H_{11} = \begin{bmatrix} \Delta_{\mathrm{CT}} - 2u_p + 5j_H^{(p)} & 0 \\ 0 & \Delta_{\mathrm{CT}} - 2u_p + 5j_H^{(p)} \end{bmatrix}, \quad H_{22} = 2\Delta_{\mathrm{CT}} - 3u_p + 9j_H^{(p)}, \tag{B23}$$

into Eq. (B4), we have

$$H_{\mathrm{eff}} = -\frac{2(t^{dp})^2}{\left(\Delta_{\mathrm{CT}} - 2u_p + 5j_H^{(p)}\right)^2} - \frac{(t^{dp})^2}{\left(\Delta_{\mathrm{CT}} - 2u_p + 5j_H^{(p)}\right)^2}\frac{4(t^{dp})^2}{2\Delta_{\mathrm{CT}} - 3u_p + 9j_H^{(p)}}. \tag{B24}$$

For $\Delta_{\mathrm{CT}} \ll 0$, we rearrange the subspace as

$$\left\{ \frac{1}{\sqrt{2}}(|\uparrow,\downarrow;\uparrow,0;0,\downarrow;\uparrow,\downarrow\rangle - |\uparrow,\downarrow;0,\downarrow;\uparrow,0;\uparrow,\downarrow\rangle),\ \frac{1}{\sqrt{2}}(|\uparrow,0;\uparrow,\downarrow;0,\downarrow;\uparrow,\downarrow\rangle - |0,\downarrow;\uparrow,\downarrow;\uparrow,0;\uparrow,\downarrow\rangle), \right.$$
$$\left. \frac{1}{\sqrt{2}}(|\uparrow,\downarrow;\uparrow,0;\uparrow,\downarrow;0,\downarrow\rangle - |\uparrow,\downarrow;0,\downarrow;\uparrow,0;\uparrow,\downarrow\rangle),\ \frac{1}{\sqrt{2}}(|\uparrow,0;\uparrow,\downarrow;\uparrow,\downarrow;0,\downarrow\rangle - |0,\downarrow;\uparrow,\downarrow;\uparrow,\downarrow;\uparrow,0\rangle) \right\},$$

and the corresponding configuration interaction matrix is

$$\begin{bmatrix} 2\Delta_{\mathrm{CT}} - 3u_p + 9j_H^{(p)} & t^{dp} & t^{dp} & 0 \\ t^{dp} & \Delta_{\mathrm{CT}} - 2u_p + 5j_H^{(p)} & 0 & t^{dp} \\ t^{dp} & 0 & \Delta_{\mathrm{CT}} - 2u_p + 5j_H^{(p)} & t^{dp} \\ 0 & t^{dp} & t^{dp} & 0 \end{bmatrix}. \tag{B25}$$

Here we are interested in $\varepsilon \sim (2\Delta_{\mathrm{CT}} - 3u_p + 9j_H^{(p)})$ and put

$$H_{00} = 2\Delta_{\mathrm{CT}} - 3u_p + 9j_H^{(p)}, \quad H_{11} = \begin{bmatrix} \Delta_{\mathrm{CT}} - 2u_p + 5j_H^{(p)} & 0 \\ 0 & \Delta_{\mathrm{CT}} - 2u_p + 5j_H^{(p)} \end{bmatrix}, \quad H_{22} = 0, \tag{B26}$$

into Eq. (B4), we have

$$H_{\mathrm{eff}} = \left(2\Delta_{\mathrm{CT}} - 3u_p + 9j_H^{(p)}\right) + \frac{2(t^{dp})^2}{\Delta_{\mathrm{CT}} - u_p + 4j_H^{(p)}} + \frac{(t^{dp})^2}{\left(\Delta_{\mathrm{CT}} - u_p + 4j_H^{(p)}\right)^2} \frac{4(t^{dp})^2}{2\Delta_{\mathrm{CT}} - 3u_p + 9j_H^{(p)}}. \tag{B27}$$

b. S = 1 sector

There are 18 states in the S = 1 sector,

$$\Big\{|0,\downarrow;\uparrow,\downarrow;\uparrow,\downarrow;0,\downarrow\rangle,\ |0,\downarrow;\uparrow,\downarrow;0,\downarrow;\uparrow,\downarrow\rangle,\ |\uparrow,\downarrow;0,\downarrow;\uparrow,\downarrow;0,\downarrow\rangle,\ |\uparrow,\downarrow;0,\downarrow;0,\downarrow;\uparrow,\downarrow\rangle,\ |\uparrow,\downarrow;\uparrow,\downarrow;0,\downarrow;0,\downarrow\rangle,$$
$$|0,\downarrow;0,\downarrow;\uparrow,\downarrow;\uparrow,\downarrow\rangle,\ \frac{1}{\sqrt{2}}(|\uparrow,0;\uparrow,\downarrow;\uparrow,\downarrow;0,\downarrow\rangle + |0,\downarrow;\uparrow,\downarrow;\uparrow,\downarrow;\uparrow,0\rangle),$$
$$\frac{1}{\sqrt{2}}(|\uparrow,0;\uparrow,\downarrow;0,\downarrow;\uparrow,\downarrow\rangle + |0,\downarrow;\uparrow,\downarrow;\uparrow,0;\uparrow,\downarrow\rangle),\ \frac{1}{\sqrt{2}}(|\uparrow,\downarrow;\uparrow,0;\uparrow,\downarrow;0,\downarrow\rangle + |\uparrow,\downarrow;0,\downarrow;\uparrow,\downarrow;\uparrow,0\rangle),$$
$$\frac{1}{\sqrt{2}}(|\uparrow,\downarrow;\uparrow,0;0,\downarrow;\uparrow,\downarrow\rangle + |\uparrow,\downarrow;0,\downarrow;\uparrow,0;\uparrow,\downarrow\rangle),\ \frac{1}{\sqrt{2}}(|\uparrow,\downarrow;\uparrow,\downarrow;\uparrow,0;0,\downarrow\rangle + |\uparrow,\downarrow;\uparrow,\downarrow;0,\downarrow;\uparrow,0\rangle),$$
$$\frac{1}{\sqrt{2}}(|\uparrow,0;0,\downarrow;\uparrow,\downarrow;\uparrow,\downarrow\rangle + |0,\downarrow;\uparrow,0;\uparrow,\downarrow;\uparrow,\downarrow\rangle),\ |\uparrow,0;\uparrow,\downarrow;\uparrow,\downarrow;\uparrow,0\rangle,\ |\uparrow,0;\uparrow,\downarrow;\uparrow,0;\uparrow,\downarrow\rangle,\ |\uparrow,\downarrow;\uparrow,0;\uparrow,\downarrow;\uparrow,0\rangle,$$
$$|\uparrow,\downarrow;\uparrow,0;\uparrow,0;\uparrow,\downarrow\rangle,\ |\uparrow,\downarrow;\uparrow,\downarrow;\uparrow,0;\uparrow,0\rangle,\ |\uparrow,0;\uparrow,0;\uparrow,\downarrow;\uparrow,\downarrow\rangle\Big\}.$$

Notice that the first six states, the second six states and the last six states are not connected via t^{dp}, in other words, we can work on one subspace. Here we will work on the last six states. What is more, notice that the first four states are also not connected with the last two via t^{dp}, for simplicity, we will work on the subspace spanned by the first four,

$$\{|\uparrow,0;\uparrow,\downarrow;\uparrow,\downarrow;\uparrow,0\rangle,\ |\uparrow,0;\uparrow,\downarrow;\uparrow,0;\uparrow,\downarrow\rangle,\ |\uparrow,\downarrow;\uparrow,0;\uparrow,\downarrow;\uparrow,0\rangle,\ |\uparrow,\downarrow;\uparrow,0;\uparrow,0;\uparrow,\downarrow\rangle\}.$$

With $\Delta_{\mathrm{CT}} \gg 0$, the configuration interaction matrix for such a subspace is

$$\begin{bmatrix} 0 & t^{dp} & t^{dp} & 0 \\ t^{dp} & \Delta_{\mathrm{CT}} - 2u_p + 5j_H^{(p)} & 0 & t^{dp} \\ t^{dp} & 0 & \Delta_{\mathrm{CT}} - 2u_p + 5j_H^{(p)} & t^{dp} \\ 0 & t^{dp} & t^{dp} & 2\Delta_{\mathrm{CT}} - 3u_p + 7j_H^{(p)} \end{bmatrix}. \tag{B28}$$

Here we are interested in $\varepsilon \sim 0$ and put

$$H_{00} = 0, \quad H_{11} = \begin{bmatrix} \Delta_{\mathrm{CT}} - 2u_p + 5j_H^{(p)} & 0 \\ 0 & \Delta_{\mathrm{CT}} - 2u_p + 5j_H^{(p)} \end{bmatrix}, \quad H_{22} = 2\Delta_{\mathrm{CT}} - 3u_p + 7j_H^{(p)}, \tag{B29}$$

into Eq. (B4), we have

$$H_{\mathrm{eff}} = -\frac{2(t^{dp})^2}{\left(\Delta_{\mathrm{CT}} - 2u_p + 5j_H^{(p)}\right)^2} - \frac{(t^{dp})^2}{\left(\Delta_{\mathrm{CT}} - 2u_p + 5j_H^{(p)}\right)^2} \frac{4(t^{dp})^2}{2\Delta_{\mathrm{CT}} - 3u_p + 7j_H^{(p)}}. \tag{B30}$$

Combining Eqs. (B24) and (B30), J is calculated as

$$J = -\frac{4(t^{dp})^2}{\left(\Delta_{\mathrm{CT}} - 2u_p + 5j_H^{(p)}\right)^2} \left[\frac{(t^{dp})^2}{\left(2\Delta_{\mathrm{CT}} - 3u_p + 8j_H^{(p)}\right) - j_H^{(p)}} - \frac{(t^{dp})^2}{\left(2\Delta_{\mathrm{CT}} - 3u_p + 8j_H^{(p)}\right) + j_H^{(p)}} \right]. \tag{B31}$$

With $\Delta_{\mathrm{CT}} \ll 0$, the basis is rearranged as

$$\{|\uparrow,\downarrow;\uparrow,0;\uparrow,0;\uparrow,\downarrow\rangle,\ |\uparrow,0;\uparrow,\downarrow;\uparrow,0;\uparrow,\downarrow\rangle,\ |\uparrow,\downarrow;\uparrow,0;\uparrow,\downarrow;\uparrow,0\rangle,\ |\uparrow,0;\uparrow,\downarrow;\uparrow,\downarrow;\uparrow,0\rangle\},$$

and the configuration interaction matrix now becomes

$$\begin{bmatrix} 2\Delta_{\mathrm{CT}}-3u_p+7j_H^{(p)} & t^{dp} & t^{dp} & 0 \\ t^{dp} & \Delta_{\mathrm{CT}}-2u_p+5j_H^{(p)} & 0 & t^{dp} \\ t^{dp} & 0 & \Delta_{\mathrm{CT}}-2u_p+5j_H^{(p)} & t^{dp} \\ 0 & t^{dp} & t^{dp} & 0 \end{bmatrix}.$$

Here we are interested in $\varepsilon \sim (2\Delta_{\mathrm{CT}}-3u_p+7j_H^{(p)})$ and put

$$H_{00}=2\Delta_{\mathrm{CT}}-3u_p+7j_H^{(p)},\quad H_{11}=\begin{bmatrix} \Delta_{\mathrm{CT}}-2u_p+5j_H^{(p)} & 0 \\ 0 & \Delta_{\mathrm{CT}}-2u_p+5j_H^{(p)} \end{bmatrix},\quad H_{22}=0, \tag{B32}$$

into Eq. (B4), we have

$$H_{\mathrm{eff}}=\left(2\Delta_{\mathrm{CT}}-3u_p+7j_H^{(p)}\right)+\frac{2(t^{dp})^2}{\Delta_{\mathrm{CT}}-u_p+2j_H^{(p)}}+\frac{(t^{dp})^2}{\left(\Delta_{\mathrm{CT}}-u_p+2j_H^{(p)}\right)^2}\frac{4(t^{dp})^2}{2\Delta_{\mathrm{CT}}-3u_p+7j_H^{(p)}}. \tag{B33}$$

Combining Eqs. (B27) and (B33), we have

$$J=-2j_H^{(p)} \tag{B34}$$

where higher orders are ignored.

3. Further reduction for $\Delta_{\mathrm{CT}} \gg 0$

For $\Delta_{\mathrm{CT}} \gg 0$, p is much deeper than d orbital, under such circumstances, p orbital can be treated as uncorrelated one with $u_p=0$ and $j_H^{(p)}=0$ [but the last $j_H^{(p)}$ term of Eq. (3) is kept]. So $\Delta_{\mathrm{CT}}=\varepsilon_d-\varepsilon_p+u_d=\varepsilon_{dp}+u_d$ and Eqs. (B17) and (B31) are simplified to

$$\begin{aligned} J&=\frac{4(t^{dp})^2}{(\varepsilon_{dp}+u_d)^2}\left(\frac{1}{u_d}+\frac{1}{\varepsilon_{dp}+u_d}\right), \\ J&=-\frac{4(t^{dp})^2}{(\varepsilon_{dp}+u_d)^2}\left(\frac{1}{2(\varepsilon_{dp}+u_d)-j_H^{(p)}}\right. \\ &\quad\left.-\frac{1}{2(\varepsilon_{dp}+u_d)+j_H^{(p)}}\right), \end{aligned} \tag{B35}$$

in accordance with the results of Koch [93].

APPENDIX C: MULTIBAND EXTENSION

For simplicity, here we consider the case $t^d=0$, $t^p=0$. In the PM phase, p orbital lies ε_{dp} below d orbital as shown in Fig. 10(a). Suppose now both d and p subsystem are FM because of strong interaction, there are two possible configurations for the whole system: FM_d-FM_p-FM in Fig. 10(b) and FM_d-FM_p-AFM in Fig. 10(c). For FM_d-FM_p-AFM, since p orbital has opposite polarization with respect to d orbital, the spin up electron will have higher energy than spin down electron. This is the difference between Fig. 10(b) and Fig. 10(c). In the following, we will study the two cases: (1) d is half filled and (2) d is empty. Since the two p orbitals are degenerate as well as the two d orbitals, in the following, we will only consider one p and one d orbital.

1. d orbital is half filled

Now we include t^{dp}. By fixing N_e to 3, the energy of FM_d-FM_p-FM and FM_d-FM_p-AFM is the sum of the lowest three energy level. For FM_d-FM_p-FM, we have

$$\begin{aligned} &E(\mathrm{FM_d}-\mathrm{FM_p}-\mathrm{FM}) \\ &\quad=-[\varepsilon_{dp}-(X_d-X_p)] \\ &\quad\quad-\frac{[\varepsilon_{dp}+(X_d-X_p)]+\sqrt{[\varepsilon_{dp}+(X_d-X_p)]^2+4(t^{dp})^2}}{2}, \end{aligned} \tag{C1}$$

where X_p, X_d is half of the Zeeman splitting of p and d orbital due to the intrinsic magnetic order as shown in Figs. 10(b) and

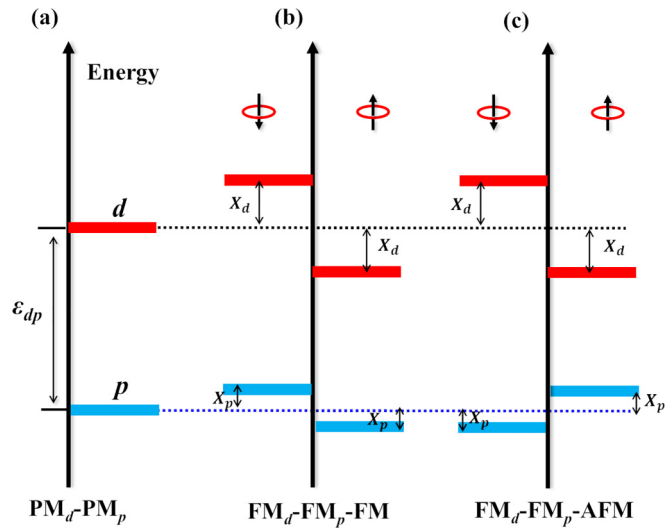

FIG. 10. (a) Energy level in the PM phase. (b) Energy level in the $\mathrm{FM_d}-\mathrm{FM_p}-\mathrm{FM}$ phase. (c) Energy level in the $\mathrm{FM_d}-\mathrm{FM_p}-$ AFM phase.

10(c). On the other hand, the energy for FM_d-FM_p-AFM is

$$\begin{aligned} &E(\mathrm{FM}_d-\mathrm{FM}_p-\mathrm{AFM}) \\ &\quad=-[\varepsilon_{dp}-(X_d+X_p)] \\ &\quad-\frac{[\varepsilon_{dp}+(X_d+X_p)]+\sqrt{[\varepsilon_{dp}+(X_d+X_p)]^2+4(t^{dp})^2}}{2}. \end{aligned} \tag{C2}$$

The energy difference between FM_d-FM_p-FM and FM_d-FM_p-AFM can be seen as a function of X_p,

$$\begin{aligned} f_1(X_p)&=E(\mathrm{FM}_d-\mathrm{FM}_p-\mathrm{FM})-E(\mathrm{FM}_d-\mathrm{FM}_p-\mathrm{AFM}) \\ &=-X_p-\frac{\sqrt{[\varepsilon_{dp}+(X_d-X_p)]^2+4(t^{dp})^2}}{2} \\ &\quad+\frac{\sqrt{[\varepsilon_{dp}+(X_d+X_p)]^2+4(t^{dp})^2}}{2}. \end{aligned} \tag{C3}$$

When $X_p=0$, $f_1(X_p)=0$ as expected. Taking derivative on X_p, we have

$$\begin{aligned} \frac{d}{dX_p}f_1(X_p)&=-1+\frac{1}{2}\frac{1}{\sqrt{1+\left(\frac{2t^{dp}}{\varepsilon_{dp}+(X_d-X_p)}\right)^2}} \\ &\quad+\frac{1}{2}\frac{1}{\sqrt{1+\left(\frac{2t^{dp}}{\varepsilon_{dp}+(X_d+X_p)}\right)^2}}<-1+\frac{1}{2}+\frac{1}{2}=0. \end{aligned} \tag{C4}$$

Combing with the fact that $f_1(X_p=0)=0$, we have $E(FM_d-FM_p-FM)<E(FM_d-FM_p-AFM)$ when $X_p\neq 0$. This is the main reason why the ground state for Δ_{CT} closing to 0 is FM_d-FM_p-FM in Fig. 4(c), not FM_d-FM_p-AFM.

2. *d* orbital is empty

Now we consider the case when d is empty, in other words, we have $N_e=2$. The energy of FM_d-FM_p-FM and FM_d-FM_p-AFM is the sum of the lowest two energy level. For FM_d-FM_p-FM, we have

$$\begin{aligned} E(\mathrm{FM}_d-\mathrm{FM}_p-\mathrm{FM})&=\frac{-[\varepsilon_{dp}-(X_d-X_p)]-\sqrt{[\varepsilon_{dp}-(X_d-X_p)]^2+4(t^{dp})^2}}{2} \\ &\quad+\frac{-[\varepsilon_{dp}+(X_d-X_p)]-\sqrt{[\varepsilon_{dp}+(X_d-X_p)]^2+4(t^{dp})^2}}{2}. \end{aligned} \tag{C5}$$

The energy for FM_d-FM_p-AFM is

$$\begin{aligned} E(\mathrm{FM}_d-\mathrm{FM}_p-\mathrm{AFM})&=\frac{-[\varepsilon_{dp}-(X_d+X_p)]-\sqrt{[\varepsilon_{dp}-(X_d+X_p)]^2+4(t^{dp})^2}}{2} \\ &\quad+\frac{-[\varepsilon_{dp}+(X_d+X_p)]-\sqrt{[\varepsilon_{dp}+(X_d+X_p)]^2+4(t^{dp})^2}}{2}. \end{aligned} \tag{C6}$$

Then $f_2(X_p)$ is given by

$$\begin{aligned} f_2(X_p)&=E(\mathrm{FM}_d-\mathrm{FM}_p-\mathrm{FM})-E(\mathrm{FM}_d-\mathrm{FM}_p-\mathrm{AFM}) \\ &=\frac{\sqrt{[\varepsilon_{dp}-(X_d+X_p)]^2+4(t^{dp})^2}}{2}+\frac{\sqrt{[\varepsilon_{dp}+(X_d+X_p)]^2+4(t^{dp})^2}}{2} \\ &\quad-\frac{\sqrt{[\varepsilon_{dp}-(X_d-X_p)]^2+4(t^{dp})^2}}{2}-\frac{\sqrt{[\varepsilon_{dp}+(X_d-X_p)]^2+4(t^{dp})^2}}{2}. \end{aligned} \tag{C7}$$

When $X_p=0$, $f_2(X_p)=0$ as expected. Taking derivative on X_p, we have

$$\frac{d}{dX_p}f_2(X_p)=\frac{1}{2}\left(\frac{1}{\sqrt{1+\left(\frac{2t^{dp}}{\varepsilon_{dp}+(X_d+X_p)}\right)^2}}-\frac{1}{\sqrt{1+\left(\frac{2t^{dp}}{\varepsilon_{dp}+(X_d-X_p)}\right)^2}}\right)+\frac{1}{2}\left(\frac{1}{\sqrt{1+\left(\frac{2t^{dp}}{\varepsilon_{dp}-(X_d-X_p)}\right)^2}}-\frac{1}{\sqrt{1+\left(\frac{2t^{dp}}{\varepsilon_{dp}-(X_d+X_p)}\right)^2}}\right). \tag{C8}$$

Since we have

$$\begin{aligned} &\varepsilon_{dp}+(X_d+X_p)>\varepsilon_{dp}+(X_d-X_p), \\ &\varepsilon_{dp}-(X_d-X_p)>\varepsilon_{dp}-(X_d+X_p), \end{aligned} \tag{C9}$$

we have $\frac{d}{dX_p}f_2(X_p)>0$. Together with the fact that $f_2(X_p=0)=0$, we have $E(\mathrm{FM}_d-\mathrm{FM}_p-\mathrm{FM})>E(\mathrm{FM}_d-\mathrm{FM}_p-\mathrm{AFM})$ when $X_p\neq 0$. Therefore, FM_d-FM_p-AFM will be the preferred ground state for Δ_{CT} closing to 0 when d is empty.

3. One *d* orbital is half-filled and one *d* orbital is empty

We now stack the above two systems together (there is no communication between these two systems). The $j_H^{(p)}$ and $j_H^{(d)}$ align the spins on the two d and two p orbitals along the same direction. Now we ask which phase is more stable, FM_d-FM_p-FM or FM_d-FM_p-AFM ? This is equivalent to calculate the sign of the following energy:

$$f(X_p)=f_1(X_p)+f_2(X_p). \tag{C10}$$

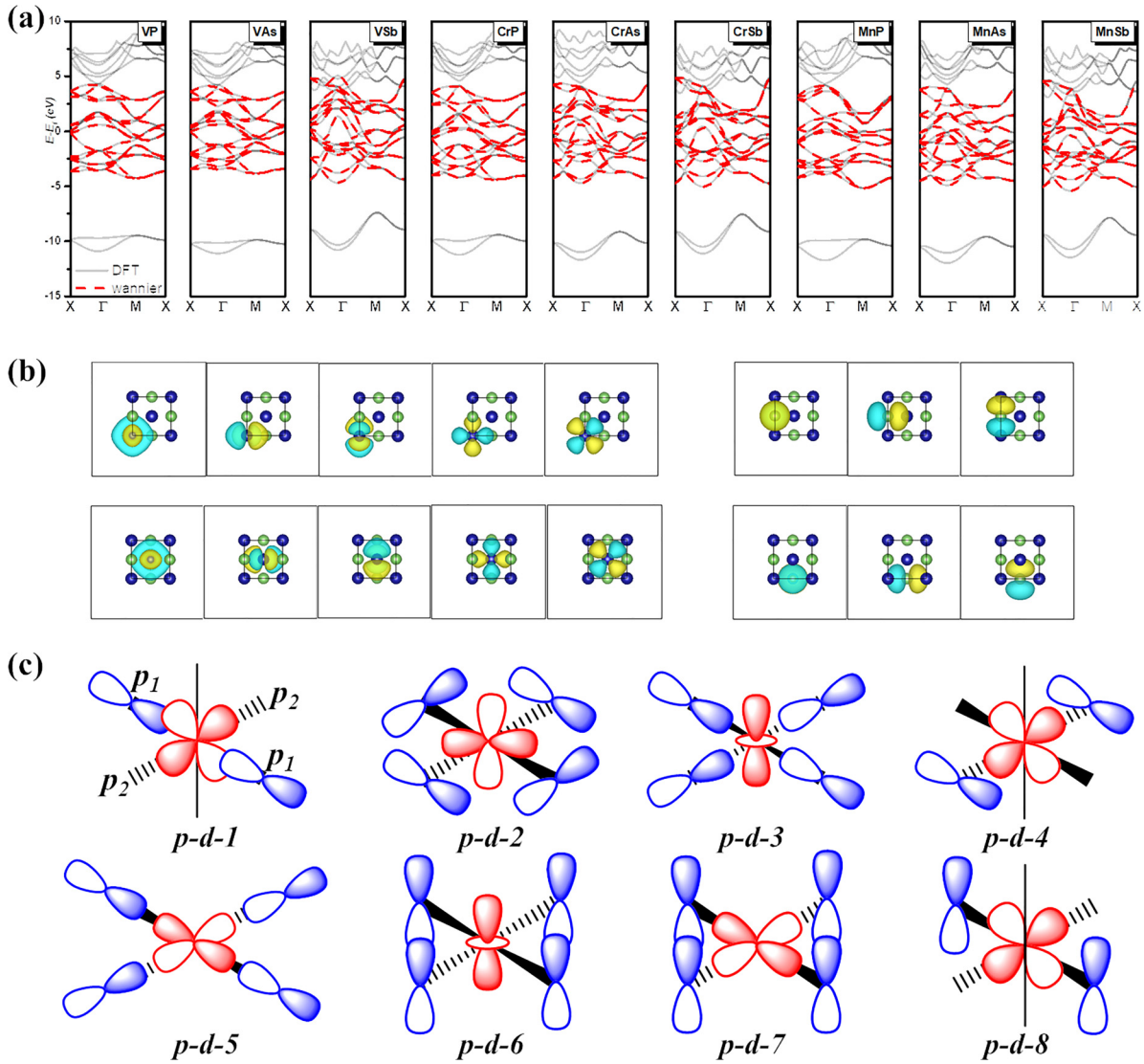

FIG. 11. (a) PBE level PM band structure of all 9 M*L* monolayers. (b) The representative maximally localized WFs of PM CrAs monolayer. (c) Eight symmetry-allowed $p-d$ hopping channel. The dashed/bold wedge represents bonds above/below the paper plane, following the Natta projection in stereochemistry. The p orbitals above (below) the paper plane is labelled as p_1 (p_2)

Obviously, we have $f(X_p = 0) = 0$. The derivative on X_p is

$$\frac{d}{dX_p} f(X_p) = -1 + \frac{1}{\sqrt{1 + \left(\frac{2t^{dp}}{\varepsilon_{dp} + (X_d - X_p)}\right)^2}} + \frac{1}{\sqrt{1 + \left(\frac{2t^{dp}}{\varepsilon_{dp} + (X_d + X_p)}\right)^2}} + \frac{1}{2}\left(\frac{1}{\sqrt{1 + \left(\frac{2t^{dp}}{\varepsilon_{dp} - (X_d - X_p)}\right)^2}} - \frac{1}{\sqrt{1 + \left(\frac{2t^{dp}}{\varepsilon_{dp} - (X_d + X_p)}\right)^2}}\right). \tag{C11}$$

As long as t^{dp} is not too large, we will have $\frac{d}{dX_p} f(X_p) > 0$. Therefore, FM$_d$-FM$_p$-AFM will be the preferred phase for partial filled d shell in general.

APPENDIX D: MODEL PARAMETERS CALCULATION

1. Single-particle part

The single-particle parameters such as on-site energy $(\varepsilon_d, \varepsilon_p)$ and hopping energy (t^{dp}, t^p, t^p) in Eq. (1) can be obtained by downfolding the full Hamiltonian into the $\{d, p\}$ subspace in Wannier90 package [94]. Explicitly, the downfolding process also allows us to obtain the following matrix element:

$$H_{\alpha\beta}(R) = \langle \phi_{0,\alpha} | \hat{H} | \phi_{R,\beta} \rangle, \tag{D1}$$

where $|\phi_{0,\alpha}\rangle$ is the maximally localized Wannier function α in home cell (index as 0) and $|\phi_{R,\beta}\rangle$ the maximally localized Wannier function (MLWF) β in cell R. When R $= 0$, $\alpha = \beta$, the above matrix element orbital energy, otherwise we obtain the hopping energy.

TABLE IV. Symmetry permitted p-d hopping channels.

Channel	d	p
$p-d-1$	d_{xz} (d_{yz})	p_{1x} (p_{2y})
$p-d-2$	d_{xy}	$p_{1y}+p_{2x}$
$p-d-3$	d_{z^2}	$p_{1x}+p_{2y}$
$p-d-4$	d_{xz} (d_{yz})	p_{2x} (p_{1y})
$p-d-5$	$d_{x^2-y^2}$	$p_{1x}+p_{2y}$
$p-d-6$	d_{z^2}	$p_{1z}+p_{2z}$
$p-d-7$	$d_{x^2-y^2}$	$p_{1z}+p_{2z}$
$p-d-8$	d_{xz} (d_{yz})	p_{1z} (p_{2z})

The PM band structure with Perdew-Burke-Ernzerhof (PBE) functional [83] and the Wannier fitted one is shown in Fig. 11(a). The corresponding 16 maximally localized WFs for CrAs monolayer are displayed in Fig. 11(b). The excellent agreement of the band structures and the small spreading of MLWFs indicates the downfolding process is quite good, which lays the foundation of cRPA calculation below. In cRPA calculation, 72 bands were used with $5\times5\times1$ k-point meshes for Brillouin zone integration. What is more, by setting all t^{dp} to zero while keeping other hopping terms, the band structure without t^{dp} can be obtained, as shown in Fig. 6(a) and Figs. 7(a) and 7(b).

The space group of this lattice structure is $P\frac{4}{n}mm$, all the $p-d$ hybridizations can be classified into eight types (labelled as $p-d-i, i=1,2,\ldots,8$) displayed in Fig. 11(c). Explicitly, the $p-d-1$ composes of d_{xz} and p_{1x} as well as the equivalent d_{yz} and p_{2y}. The p orbitals of L above and below M plane are marked as p_1 and p_2 correspondingly. And the orbital contribution to other $p-d$ channels is listed in Table IV. When there is no buckling in this structure, the point group becomes D_{4h}. Only $p-d-2$, $p-d-3$, $p-d-5$ and $p-d-8$ exist in D_{4h} with $p-d-5$ the strongest, which can be as large as 1.3 eV in cuprates [95]. When the structure is buckled, there are no 90° and 180° M-L-M angles and $p-d-1$, $p-d-4$, $p-d-6$, $p-d-7$ appear.

2. Interaction term

In solid, the Coulomb potential in solid is screened by electronic polarizability and is thus renormalized. The constrained random phase approximation (cRPA) provides a systematic first-principles technique for the construction of low-energy Hamiltonians where the interaction part is calculated [96–98].

Clearly there is a degree of freedom in choosing target subspace. Different subspace orbitals will give different interaction parameters (u, u', j_H), at the same time, the number of parameters are also changed as well as the on-site energy and hopping strength. In MHI, it is often enough to choose part of M d orbitals as the target space. In our case here, we need to treat Eq. (1), so both M d and L p orbitals are chosen to be target subspace.

After calculation, the value of the screened interaction (W^r) between local orbitals is expressed as 4-index interaction matrix [66],

$$U^{(S)}_{m_1m_2m_3m_4}(\omega)=\langle\phi_{m_1}\phi_{m_2}|W^r(\omega)|\phi_{m_3}\phi_{m_4}\rangle, \tag{D2}$$

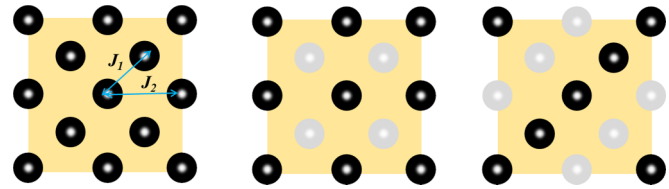

FIG. 12. Schematic view of three ordered magnetic states. The black/white solid sphere represents up/down spin magnetic moment, respectively.

where ω is the frequency, which describes the dynamical effect (only $\omega=0$ is used here), the $|\phi_m\rangle$ is the localized Wannier orbitals, and (S) is added for specifying the angular symmetry of the localized Wannier orbitals. Here we take cubic angular harmonics approximation so (S) is cubic. Most matrix elements are of the order of 0.1 eV or less, except for 2-index reduced interaction matrices [66]. Furthermore, 2-index matrix can be further simplified as scalars, which is used in model Hamiltonian like Eq. (1). For either d or p shell, there are three independent intrashell values,

$$\begin{aligned} u &= \frac{1}{N}\sum_{m=1}^{N} U^{\text{cubic}}_{\text{mmmm}},\\ u' &= \frac{1}{N(N-1)}\sum_{m\neq n}^{N} U^{\text{cubic}}_{\text{mnmn}},\\ j_H &= \frac{1}{N(N-1)}\sum_{m\neq n}^{N} U^{\text{cubic}}_{\text{mnnm}}, \end{aligned} \tag{D3}$$

where $N=5$ (3) for d (p) shell. As for the intershell Coulomb interaction, only the density-density interaction u_{dp} is calculated as all the other interaction terms are nearly zero,

$$u_{dp}=\frac{1}{15}\sum_{m,n} U^{\text{cubic}}_{\text{dmpn}}. \tag{D4}$$

APPENDIX E: EXCHANGE COUPLING STRENGTH

Here energy mapping method was applied to calculate J. In this method, the total energies of different spin configuration are calculated and the exchange interactions are fitted to the energies of different spin configuration. For example, to map J_1, J_2 out, three symmetric magnetic orders were considered in a $2\times2\times1$ supercell as shown in Fig. 12: FM with magnetic ordering momentum $\mathbf{q}$=(0, 0), checkboard AFM (c-AFM) with $\mathbf{q}$=(π, π) and strip AFM (s-AFM) with $\mathbf{q}=(0,\pi)$. The spin model used here is the Heisenberg model,

$$H=\sum_{i,j} J_{ij}\vec{S}_i\cdot\vec{S}_j, \tag{E1}$$

where summation is over J_1 and J_2 as defined in Fig. 12. And the energy for these three magnetic orders are calculated

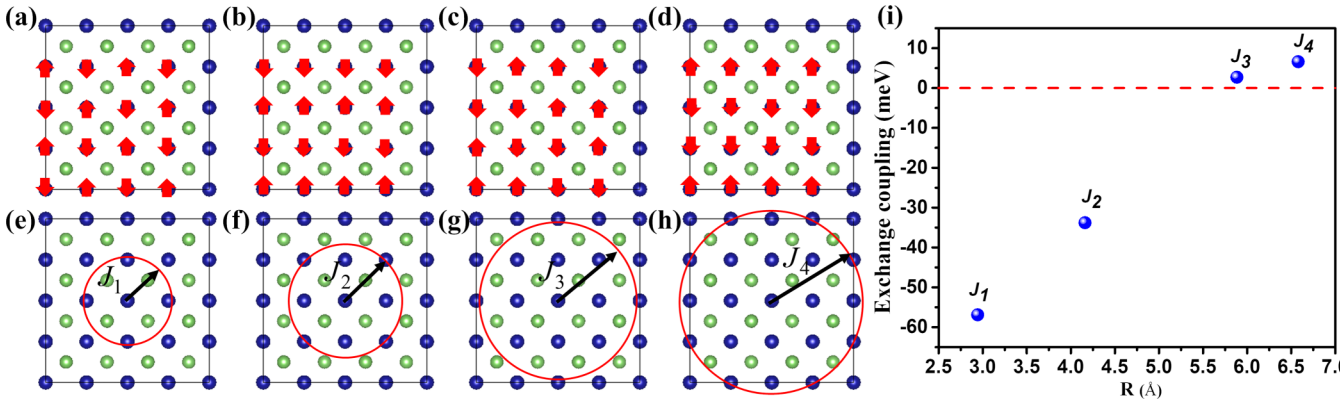

FIG. 13. (a) Checkboard AFM configuration. (b) Strip AFM configuration. (c) Bi-strip AFM configuration. (d) Antistripe AFM configuration. The up and down red arrow in (a)–(d) represents the orientation of local magnetic moments. Real space illustration of range of (e) J_1, (f) J_2, (g) J_3, and (h) J_4. (i) Strength of J with respect to the interaction distance.

as

$$E_{\mathrm{FM}} = E_0 + 16J_1S^2 + 16J_2S^2.$$
$$E_{\mathrm{c-AFM}} = E_0 - 16J_1S^2 + 16J_2S^2.$$
$$E_{\mathrm{s-AFM}} = E_0 - 16J_2S^2. \quad \text{(E2)}$$

where E_0 is a reference energy. These energies are obtained via HSE06 functional. By taking E_{FM}, $E_{\mathrm{c-AFM}}$ and $E_{\mathrm{s-AFM}}$ into above equations, J_1 and J_2 can be obtained for a given S.

When S is in the classical limit, energy mapping method has shown success in FM CHI and MHI as S is easy to define there [67,99]. This is not the case in FNCTEI as L is also polarized. However the net magnetic moment on L is quite small [see Fig. 9(g)], which makes it inappropriate to denote an integer magnetic moment attached to L. What is more, to compensate the holes on As, extra electrons are back donated to Cr, so the magnetic moment on Cr is a slightly larger than that without back-donation. So here we make the following simplification: by treating a Cr and its nearby four As as a whole, it is possible to denote an half-integer $S = 3/2$ to it. In this way, the J can be evaluated by energy mapping method.

The calculated relative energy of E_{FM}, $E_{\mathrm{c-AFM}}$ and $E_{\mathrm{s-AFM}}$ for CrAs monolayer is 0, +0.511 and +0.632 eV/Cr, respectively. By taking $S = 3/2$, the obtained J_1 and J_2 in CrAs monolayer is –56.8 and –39.7 meV.

APPENDIX F: SCALING OF EXCHANGE INTERACTION STRENGTH WITH RESPECT TO DISTANCE

Since the spreading of MMO provides a natural length scale, it is interesting to see whether J have different behavior inside and outside this length scale. To see this, we use energy mapping method on a 4×4 supercell and calculate J up to J_4 as shown in Figs. 13(a)–13(h). As it is difficult for HSE06 to handle such a large supercell, here we used SCAN functional for a compromise. The energies of the different magnetic orders are

$$E_{\mathrm{FM}} = E_0 + 32J_1S^2 + 32J_2S^2 + 32J_3S^2 + 64J_4S^2,$$
$$E_{\mathrm{c-AFM}} = E_0 - 32J_1S^2 + 32J_2S^2 + 32J_3S^2 - 64J_4S^2,$$
$$E_{\mathrm{s-AFM}} = E_0 - 32J_2S^2 + 32J_3S^2,$$
$$E_{\mathrm{bis-AFM}} = E_0 - 32J_3S^2,$$
$$E_{\mathrm{antis-AFM}} = E_0 + 16J_1S^2 - 32J_4S^2. \quad \text{(F1)}$$

The result is displayed in Fig. 13(i). It is clear that J experiences a sharp decrease from J_2 to J_3. From this result, it is reasonable to use J_1 and J_2 obtained in Appendix E in evaluating T_c in the main text. For J_1 and J_2, their interaction lengths are within the spreading of MMOs, while for J_3 and

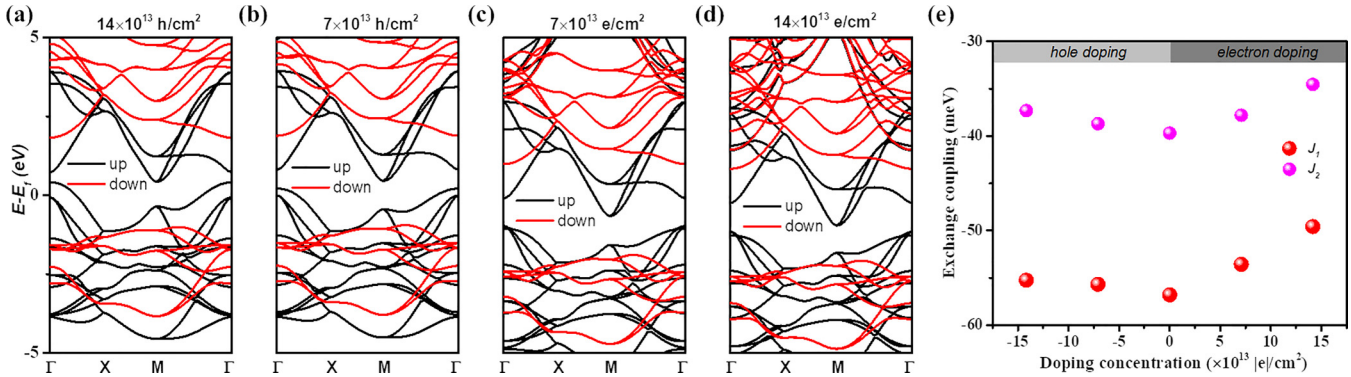

FIG. 14. [(a)–(d)] Band structures for different doping concentration. (e) Evolution of J_1 and J_2 with respect to doping. Data of pristine CrAs is also shown for comparison.

J_4, they are beyond the spreading of MMOs. Therefore, MMO indeed provides a natural length scale for J.

APPENDIX G: DOPING DEPENDENCE OF EXCHANGE INTERACTION STRENGTH

Here hole and electron doping with concentration at 7×10^{13} h(e)/cm^2 [corresponds to 0.125 h(e)/unit cell] and 14×10^{13} h(e)/cm^2 [0.25 h(e)/unit cell] were considered in FM CrAs monolayer. Such doping concentration corresponds to moderate to heavy doping for 2D materials. The results are displayed in Fig. 14. With hole doping, the Fermi level cuts the Fermi pocket at Γ point as shown in Figs. 14(a) and 14(b). With electron doping, the Fermi level cuts the Fermi pocket at M point as shown in Figs. 14(c) and 14(d). Figure 14(e) shows the dependence of J_1 and J_2 on doping. It is clear both J_1 and J_2 decrease with either hole or electron doping. Therefore, the FM order is destabilized when the electron filling deviates from the ideal filling.

[1] M. Brando, D. Belitz, F. M. Grosche, and T. R. Kirkpatrick, Rev. Mod. Phys. **88**, 025006 (2016).
[2] K. Hasegawa, M. Isobe, T. Yamauchi, H. Ueda, J.-I. Yamaura, H. Gotou, T. Yagi, H. Sato, and Y. Ueda, Phys. Rev. Lett. **103**, 146403 (2009).
[3] D. Meng, H. Guo, Z. Cui, C. Ma, J. Zhao, J. Lu, H. Xu, Z. Wang, X. Hu, Z. Fu *et al.*, Proc. Natl. Acad. Sci. USA **115**, 2873 (2018).
[4] R. Chen, F. Luo, Y. Liu, Y. Song, Y. Dong, S. Wu, J. Gao, F. Yang, A. N'Diaye, P. Shafer *et al.*, Nat. Commun. **12**, 3952 (2021).
[5] Q. Jin, Z. Wang, Q. Zhang, Y. Yu, S. Lin, S. Chen, M. Qi, H. Bai, A. Huon, Q. Li, L. Wang, X. Yin, C. S. Tang, A. T. S. Wee, F. Meng, J. Zhao, J.-O. Wang, H. Guo, C. Ge, C. Wang *et al.* Phys. Rev. Lett.**,128**, 017202 (2022).
[6] Y. Nagaoka, Phys. Rev. **147**, 392 (1966).
[7] H. Tasaki, Phys. Rev. B **40**, 9192 (1989).
[8] A. Mielke, J. Phys. A: Math. Gen. **24**, L73 (1991).
[9] A. Mielke, J. Phys. A: Math. Gen. **24**, 3311 (1991).
[10] H. Tasaki, Phys. Rev. Lett. **69**, 1608 (1992).
[11] A. Mielke and H. Tasaki, Commun. Math. Phys. **158**, 341 (1993).
[12] I. Hase, T. Yanagisawa, Y. Aiura, and K. Kawashima, Phys. Rev. Lett. **120**, 196401 (2018).
[13] H. Wang, Z. Chen, and Z. Liu, Appl. Surf. Sci. **585**, 152536 (2022).
[14] S. Q. Shen, Phys. Rev. B **57**, 6474 (1998).
[15] Y. Li, E. H. Lieb, and C. Wu, Phys. Rev. Lett. **112**, 217201 (2014).
[16] J. K. Furdyna, J. Appl. Phys. **64**, R29 (1988).
[17] H. Ohno, A. Shen, F. Matsukura, A. Oiwa, A. Endo, S. Katsumoto, and Y. Iye, Appl. Phys. Lett. **69**, 363 (1996).
[18] T. Dietl, H. Ohno, F. Matsukura, J. Cibert, and D. Ferrand, Science **287**, 1019 (2000).
[19] J. M. D. Coey, M. Venkatesan, and C. B. Fitzgerald, Nat. Mater. **4**, 173 (2005).
[20] K. Zhao, Z. Deng, X. C. Wang, W. Han, J. L. Zhu, X. Li, Q. Q. Liu, R. C. Yu, T. Goko, B. Frandsen *et al.*, Nat. Commun. **4**, 1442 (2013).
[21] T. Dietl and H. Ohno, Rev. Mod. Phys. **86**, 187 (2014).
[22] H. A. Kramers, Physica **1**, 182 (1934).
[23] P. W. Anderson, Phys. Rev. **79**, 350 (1950).
[24] J. B. Goodenough, Phys. Rev. **100**, 564 (1955).
[25] J. B. Goodenough, J. Phys. Chem. Solids **6**, 287 (1958).
[26] J. Kanamori, J. Phys. Chem. Solids **10**, 87 (1959).
[27] P. W. Anderson, Phys. Rev. **115**, 2 (1959).
[28] J. Zaanen, G. A. Sawatzky, and J. W. Allen, Phys. Rev. Lett. **55**, 418 (1985).
[29] J. Zaanen and G. A. Sawatzky, J. Solid State Chem. **88**, 8 (1990).
[30] D. I. Khomskii, *Transition Metal Compounds* (Cambridge University Press, Cambridge, 2014).
[31] P. Schwob and O. Vogt, Phys. Lett. A **24**, 242 (1967).
[32] B. T. Matthias, R. M. Bozorth, and J. H. Van Vleck, Phys. Rev. Lett. **7**, 160 (1961).
[33] B. Huang, G. Clark, E. Navarro-Moratalla, D. R. Klein, R. Cheng, K. L. Seyler, D. Zhong, E. Schmidgall, M. A. McGuire, D. H. Cobden *et al.*, Nature (London) **546**, 270 (2017).
[34] C. Gong, L. Li, Z. Li, H. Ji, A. Stern, Y. Xia, T. Cao, W. Bao, C. Wang, Y. Wang *et al.*, Nature (London) **546**, 265 (2017).
[35] C. Huang, J. Feng, J. Zhou, H. Xiang, K. Deng, and E. Kan, J. Am. Chem. Soc. **141**, 12413 (2019).
[36] A. Fert, Rev. Mod. Phys. **80**, 1517 (2008).
[37] A. Avsar, H. Ochoa, F. Guinea, B. Özyilmaz, B. J. van Wees, and I. J. Vera-Marun, Rev. Mod. Phys. **92**, 021003 (2020).
[38] X. Li and J. Yang, Natl. Sci. Rev. **3**, 365 (2016).
[39] M. A. Korotin, V. I. Anisimov, D. I. Khomskii, and G. A. Sawatzky, Phys. Rev. Lett. **80**, 4305 (1998).
[40] J. Kuneš, V. Křápek, N. Parragh, G. Sangiovanni, A. Toschi, and A. V. Kozhevnikov, Phys. Rev. Lett. **109**, 117206 (2012).
[41] T. Mizokawa, H. Namatame, A. Fujimori, K. Akeyama, H. Kondoh, H. Kuroda, and N. Kosugi, Phys. Rev. Lett. **67**, 1638 (1991).
[42] V. Bisogni, S. Catalano, R. J. Green, M. Gibert, R. Scherwitzl, Y. Huang, V. N. Strocov, P. Zubko, S. Balandeh, J.-M. Triscone *et al.*, Nat. Commun. **7**, 13017 (2016).
[43] D. I. Khomskii and G. A. Sawatzky, Solid State Commun. **102**, 87 (1997).
[44] X. Li and J. Yang, J. Am. Chem. Soc. **141**, 109 (2019).
[45] T. Mizokawa, D. I. Khomskii, and G. A. Sawatzky, Phys. Rev. B **61**, 11263 (2000).
[46] V. J. Emery, Phys. Rev. Lett. **58**, 2794 (1987).
[47] P. Hansmann, N. Parragh, A. Toschi, G. Sangiovanni, and K. Held, New J. Phys. **16**, 033009 (2014).
[48] R. Eder, J. van den Brink, and G. A. Sawatzky, Phys. Rev. B **54**, R732 (1996).
[49] Y. Claveau, B. Arnaud, and S. Di Matteo, Eur. J. Phys. **35**, 035023 (2014).
[50] S. Nimkar, D. D. Sarma, and H. R. Krishnamurthy, Phys. Rev. B **47**, 10927(R) (1993).
[51] F. C. Zhang and T. M. Rice, Phys. Rev. B **37**, 3759 (1988).

[52] Z. Ban and M. Sikirica, Acta Cryst. **18**, 594 (1965).
[53] M. Shatruk, J. Solid State Chem. **272**, 198 (2019).
[54] K. A. Filsinger, W. Schnelle, P. Adler, G. H. Fecher, M. Reehuis, A. Hoser, J.-U. Hoffmann, P. Werner, M. Greenblatt, and C. Felser, Phys. Rev. B **95**, 184414 (2017).
[55] J. Gu, Z. Zhao, J. Huang, B. G. Sumpter, and Z. Chen, ACS Nano **15**, 6233 (2021).
[56] V. Johnson and W. Jeischko, J. Solid State Chem. **11**, 161 (1974).
[57] R. Pöttgen and D. Johrendt, Z. Naturforsch. B **63**, 1135 (2008).
[58] X. Li, X. Wu, and J. Yang, J. Am. Chem. Soc. **136**, 5664 (2014).
[59] S.-W. Park, H. Mizoguchi, K. Kodama, S. Shamoto, T. Otomo, S. Matsuishi, T. Kamiya, and H. Hosono, Inorg. Chem. **52**, 13363 (2013).
[60] Y. Kamihara, T. Watanabe, M. Hirano, and H. Hosono, J. Am. Chem. Soc. **130**, 3296 (2008).
[61] F.-C. Hsu, J.-Y. Luo, K.-W. Yeh, T.-K. Chen, T.-W. Huang, P. M. Wu, Y.-C. Lee, Y.-L. Huang, Y.-Y. Chu, D.-C. Yan, and M.-K. Wu, Proc. Natl. Acad. Sci. USA **105**, 14262 (2008).
[62] Q.-Y. Wang, Z. Li, W.-H. Zhang, Z.-C. Zhang, J.-S. Zhang, W. Li, H. Ding, Y.-B. Ou, P. Deng, K. Chang *et al.*, Chinese Phys. Lett. **29**, 037402 (2012).
[63] A. E. Bocquet, T. Mizokawa, T. Saitoh, H. Namatame, and A. Fujimori, Phys. Rev. B **46**, 3771 (1992).
[64] W. Witczak-Krempa, G. Chen, Y. B. Kim, and L. Balents, Annu. Rev. Condens. Matter Phys. **5**, 57 (2014).
[65] G. Cao and P. Schlottmann, Rep. Prog. Phys. **81**, 042502 (2018).
[66] A. van Roekeghem, L. Vaugier, H. Jiang, and S. Biermann, Phys. Rev. B **94**, 125147 (2016).
[67] F. Lou, X. Y. Li, J. Y. Ji, H. Y. Yu, J. S. Feng, X. G. Gong, and H. J. Xiang, J. Chem. Phys. **154**, 114103 (2021).
[68] P. C. Hohenberg, Phys. Rev. **158**, 383 (1967).
[69] N. D. Mermin and H. Wagner, Phys. Rev. Lett. **17**, 1133 (1966).
[70] M.-H. Whangbo, E. E. Gordon, H. Xiang, H.-J. Koo, and C. Lee, Acc. Chem. Res. **48**, 3080 (2015).
[71] Z. Liu, G. Zhao, B. Liu, Z. F. Wang, J. Yang, and F. Liu, Phys. Rev. Lett. **121**, 246401 (2018).
[72] J. L. Lado and J. Fernéndez-Rossier, 2D Mater. **4**, 035002 (2017).
[73] B. Wang, Y. Zhang, L. Ma, Q. Wu, Y. Guo, X. Zhang, and J. Wang, Nanoscale **11**, 4204 (2019).
[74] K. Selte, A. Kjekshus, W. E. Jamison, A. F. Andresen, and J. E. Engebretsen, Acta Chem. Scand. **25**, 1703 (1971).
[75] W. Wu, J. Cheng, K. Matsubayashi, P. Kong, F. Lin, C. Jin, N. Wang, Y. Uwatako, and J. Luo, Nat. Commun. **5**, 5508 (2014).
[76] A.-N. Ma, P.-J. Wang, and C.-W. Zhang, Nanoscale **12**, 5464 (2020).
[77] B. Anasori, M. R. Lukatskaga, and Y. Gogotsi, Nat. Rev. Mater. **2**, 16098 (2017).
[78] S. Yang, P. Zhang, A. S. Nia, and X. Feng, Adv. Mater. **32**, 1907857 (2020).
[79] W. Ding, J. Zeng, W. Qin, P. Cui, and Z. Zhang, Phys. Rev. Lett. **124**, 027002 (2020).
[80] C. Ding, G. Gong, Y. Liu, F. Zheng, Z. Zhang, H. Yang, Z. Li, Y. Xing, J. Ge, K. He *et al.*, ACS Nano **13**, 10434 (2019).
[81] T. Devakul and L. Fu, Phys. Rev. X **12**, 021031 (2022).
[82] G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11169 (1996).
[83] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[84] A. Togo and I. Tanaka, Scr. Mater. **108**, 1 (2015).
[85] G. J. Martyna, M. L. Klein, and M. Tuckerman, J. Chem. Phys. **97**, 2635 (1992).
[86] J. Heyd, G. E. Scuseria, and M. Ernzerhof, J. Chem. Phys. **118**, 8207 (2003).
[87] J. Heyd and G. E. Scuseria, J. Chem. Phys. **121**, 1187 (2004).
[88] J. Sun, A. Ruzsinszky, and J. P. Perdew, Phys. Rev. Lett. **115**, 036402 (2015).
[89] J. Sun, R. C. Remsing, Y. Zhang, Z. Sun, A. Ruzsinszky, H. Peng, Z. Yang, A. Paul, U. Waghmare, X. Wu, M. L. Klein, and J. P. Perdew, Nat. Chem. **8**, 831 (2016).
[90] A. I. Liechtenstein, V. I. Anisimov, and J. Zaanen, Phys. Rev. B **52**, 5467(R) (1995).
[91] V. I. Anisimov, F. Aryasetiawan, and A. I. Liechtenstein, J. Phys.: Condens. Matter **9**, 767 (1997).
[92] J. P. Perdew, A. Ruzsinszky, J. Tao, V. N. Staroverov, G. E. Scuseria, and G. I. Csonka, J. Chem. Phys. **123**, 062201 (2005).
[93] E. Koch, Exchange mechanisms, in *Correlated Electrons: From Models to Materials*, edited by E. Pavarini, E. Koch, F. Anders, and M. Jarrell, Lecture Notes of the Autumn School Correlated Electrons 2012 Vol. 2 (Forschungszentrum Jülich, 2012).
[94] A. A. Mostofi, J. R. Yates, Y.-S. Lee, I. Souza, D. Vanderbilt, and N. Marzari, Comput. Phys. Commun. **178**, 685 (2008).
[95] M. Ogata and H. Fukuyama, Rep. Prog. Phys. **71**, 036501 (2008).
[96] F. Aryasetiawan, M. Imada, A. Georges, G. Kotliar, S. Biermann, and A. I. Lichtenstein, Phys. Rev. B **70**, 195104 (2004).
[97] T. Miyake and F. Aryasetiawan, Phys. Rev. B **77**, 085122 (2008).
[98] L. Vaugier, H. Jiang, and S. Biermann, Phys. Rev. B **86**, 165105 (2012).
[99] X. Li and J. Yang, J. Mater. Chem. C **2**, 7071 (2014).