Supplementary Material for
"Molecular Mott states in deficient spinel GaV_4S_8"

Heung-Sik Kim,[1,2] Kristjan Haule,[1] and David Vanderbilt[1]

[1]*Department of Physics and Astronomy,*

Rutgers University, Piscataway, NJ 08854, USA

[2]*Department of Physics, Kangwon National University, Chuncheon 24341, Korea*

A. Density functional theory calculations

For unit cell optimizations (cell volume and shape) and relaxations of initial internal coordinates, the Vienna *ab-initio* Simulation Package (VASP), which employs the projector-augmented wave (PAW) basis set [1, 2], was used for density functional theory (DFT) calculations in this work. 330 eV of plane-wave energy cutoff (PREC=high) and $15\times15\times15$ Γ-centered k-grid sampling were employed. For the treatment of electron correlations within DFT, several exchange-correlation functional were employed, including Ceperley-Alder (CA) parametrization of local density approximation [3], Perdew-Burke-Ernzerhof generalized gradient approximation (PBE) [4] and its revision for crystalline solids (PBEsol) [5], SCAN meta-GGA functional [6], DFT+U [7] on top of LDA, PBE, and PBEsol, and HSE06 hybrid functional [8, 9]. 10^{-4} eV/Å of force criterion was employed for structural optimizations.

B. Cluster dynamical mean-field theory calculations

A fully charge-self-consistent dynamical mean-field method[10], implemented in DFT + Embedded DMFT (eDMFT) Functional code (http://hauleweb.rutgers.edu/tutorials/) which is combined with WIEN2K code[11], is employed for computations of electronic properties and optimizations of internal coordinates[12]. At the DFT level the Perdew-Wang (PW) local density approximation is employed, which was argued to yield the best agreement of lattice properties when combined with DMFT[13]. $15\times15\times15$ Γ-centered k-grid was used to sample the first Brillouin zone with $RK_{\mathrm{max}} = 7.0$. A force criterion of 10^{-4} Ry/Bohr was adopted for optimizations of internal coordinates. The cubic lattice parameter was fixed to be the experimental value reported in Ref. 14.

A continuous-time quantum Monte Carlo method in the hybridization-expansion limit (CT-HYB) was used to solve the auxiliary quantum impurity problem[15]. For the CT-HYB calculations, up to 3×10^{10} Monte Carlo steps were employed for each Monte Carlo run. In most runs temperature was set to be 232K, but in calculations with 8 molecular orbitals (MOs) ($T^2 \oplus E \oplus T_a^1$ in Fig. 1 in the main text) as the correlated subspace it was increased up to 1160K because of the increased computational cost. -10 to +10 eV of hybridization window (with respect to the Fermi level) was chosen, and the on-site Coulomb interaction parameters U and J_{H} for V $t_{2\mathrm{g}}$ orbitals were varied within the range of $6 \sim 8$ eV and $0 \sim 1.5$ eV, respectively. A simplified Ising-type

(density-density terms only) Coulomb interaction was employed in this work, and it was tested that the use of full Coulomb interaction yields only quantitative difference in results with MO-T^2 and $T^2 \oplus E$ (not tested for MO-$T^2 \oplus E \oplus T_a^1$ case due to the high cost, see Sec. D). A nominal double counting scheme was used, with the MO occupations for double counting corrections for the V_4 cluster were chosen to be 1 or 5, depending on the choice of correlated subspace; 1 for MO-T^2 and $T^2 \oplus T_a^1$, and 5 for other cases with including E in the correlated subspace.

In the CT-HYB calculations of the $T^2 \oplus E \oplus T_a^1$ MO subspace, MO multiplet states with the occupancy $n \leq 7$ were kept (26,333 states out of 4^8 = 65,536 states in the 8 orbital Fock space) to reduce the computational cost, where the average impurity occupancy was ~ 5. It was checked that the sum of probabilities for $n \geq 8$ configurations are less than 1 percent. The high-frequency tail of the Green's function was calculated via the Hubbard-I approximation.

We comment that, due to the quite small intercluster hybridization, the perturbation order is small in our CT-HYB formalism, with the average perturbation order being less than 80 for the case of the largest correlated subspace ($T^2 \oplus E \oplus T_a^1$). In addition, it is shown below that the cubic symmetry of the V_4 cluster enforces the form of Coulomb interaction matrix between the molecular orbitals to be identical to that of atomic orbitals, at least for the T^2 molecular orbital. Hence the negative sign problem in our CT-HYB formalism is suppressed, which greatly facilitates the computation in addition to the small perturbation order.

For the computational resources, we used 8 Intel Xeon E5-2680 v4 CPUs (2.4 GHz, total 112 CPU cores) and about 400GB of memory for the case of the largest correlated subspace ($T^2 \oplus E \oplus T_a^1$). Even with the use of a high temperature $T = 1160$K and truncation of superstates in the CT-HYB stage, employing the full Coulomb interaction or a lower temperature such as $T = 780$K requires allocation of a memory size that exceeds the limitation of our hardware (512 GB). For the continuation of this study, either larger computational resources or more efficient ways to treat the less-occupied T_a^1 orbital would be necessary.

As for possible inter-site, intra-cluster Coulomb repulsion terms, the screened Coulomb repulsion in solid state compounds can be fitted to a Yukawa-like form $V(r) \simeq e^{-\lambda r}/\epsilon r$, and by comparing the unscreened and screened Coulomb parameters (U and J), one can obtain the screening length λ and electric permeability ϵ for a given system [16]. The values of unscreened U and J can be directly computed by using the local orbital projectors, and it was shown in a recent study that reasonable values of the screened U and J for the $3d$ transition metal elements in our DFT+DMFT implementation are 10 and 1 eV respectively [17]. With these we get $\lambda \simeq 0.52$ and

$\epsilon \simeq 1.09$, which yields $\int d\mathbf{r} \int d\mathbf{r}' \rho(\mathbf{r})\rho(\mathbf{r}')\frac{e^{-\lambda|\mathbf{r}-\mathbf{r}'-\mathbf{R}|}}{\epsilon|\mathbf{r}-\mathbf{r}'-\mathbf{R}|} \simeq 0.15$ eV ($|\mathbf{R}|$ being the intra-cluster V-V distance). This value is an order of magnitude smaller than the intra-cluster hybridization and on-site Coulomb interaction, so we conclude that the effect of intra-cluster, inter-site Coulomb repulsion is insignificant.

C.   Projecting the on-site Coulomb interactions onto the MO subspace

Note that the U and J_{H} are parameters defined for the atomic orbitals, which should be unitary transformed and projected onto the MOs for the impurity solver. More generally, the Coulomb repulsion matrix elements $U_{m_1,m_2,m_1',m_2'}$ at an atomic site have the form,

$$U_{m_1,m_2,m_1',m_2'} = \sum_{m,k} \frac{4}{2\pi+1} \langle Y_{lm_1}|Y_{km}|Y_{lm_1'}\rangle \langle Y_{lm_2}|Y_{km}^*|Y_{lm_2'}\rangle F^k, \tag{1}$$

where F^k are nonzero only for $k = 0, 2, 4$ for d-orbitals ($l = 2$) and $\langle Y_{lm_1}|Y_{km}|Y_{lm_1'}\rangle$ are Clebsch-Gordan coefficients. We introduce the MO states

$$|D_\alpha\rangle = \sum_{im} (Q^\dagger)_\alpha^{im} |Y_{lm}^i\rangle, \tag{2}$$

where Q is the unitary transform between the MO and the atomic orbitals, and α and $i = 1, \cdots, 4$ are the MO orbital and atomic site indices respectively. Then the Coulomb repulsion matrix elements for the MO states $U_{\alpha_1,\alpha_2,\alpha_1',\alpha_2'}$ can be written as

$$U_{\alpha_1,\alpha_2,\alpha_1',\alpha_2'} = \sum_{i,m,k} \frac{4}{2\pi+1} \langle D_{\alpha_1}|Y_{km}^i|D_{\alpha_1'}\rangle \langle D_{\alpha_2}|Y_{km}^{i*}|D_{\alpha_2'}\rangle F^k \tag{3}$$

$$\sim (QQQ^\dagger Q^\dagger)_{\{\alpha\}}^{i\{m\}} U_{\{m\}}^i. \tag{4}$$

Note that the inter-site Coulomb interactions were ignored here, which can be considered insignificant in $3d$ transition metal compounds.

Below we show explicitly how the on-site Coulomb interactions projected onto the T^2 triplet subspace should look like. As shown in Fig. 1 in the main text, electronic structure near the Fermi level ([-1eV, 1eV] window with respect to the Fermi level) is dominated by the atomic $t_{2\mathrm{g}}$ orbitals of V due to the distorted but prevalent cubic VS$_6$ octahedral environment. Therefore choosing 12 $t_{2\mathrm{g}}$ orbitals as our main interest is a reasonable choice. For simplicity we chose the Kanamori form

of the Coulomb interaction, which is written in a normal-ordered form as follows;

$$
\begin{aligned}
\hat{H}_K = -\sum_i \Big[& (U - 2J) \sum_{mm'} \hat{d}^\dagger_{im\uparrow}\hat{d}^\dagger_{im'\downarrow}\hat{d}_{im\uparrow}\hat{d}_{im'\downarrow} \\
& + 2J \sum_m \hat{d}^\dagger_{im\uparrow}\hat{d}^\dagger_{im\downarrow}\hat{d}_{im\uparrow}\hat{d}_{im\downarrow} \\
& + \frac{U - 3J}{2} \sum_{m\neq m',\sigma} \hat{d}^\dagger_{im\sigma}\hat{d}^\dagger_{im'\sigma}\hat{d}_{im\sigma}\hat{d}_{im'\sigma} \\
& - J \sum_{m\neq m'} \hat{d}^\dagger_{im\uparrow}\hat{d}^\dagger_{im'\downarrow}\hat{d}_{im\downarrow}\hat{d}_{im'\uparrow} \\
& - J \sum_{m\neq m'} \hat{d}^\dagger_{im\uparrow}\hat{d}^\dagger_{im\downarrow}\hat{d}_{im'\downarrow}\hat{d}_{im'\uparrow} \Big].
\end{aligned}
\tag{5}
$$

Here i, σ, and m, m' are site, spin, and orbital indices for Cartesian t_{2g} orbitals ($d_{xz,yz,xy}$) respectively.

Now we introduce the MO creation/annihilation operators;

$$
\hat{d}_{im\sigma} = \sum_\alpha Q^\alpha_{im} \hat{D}_{\alpha\sigma}
\tag{6}
$$

$$
\hat{d}^\dagger_{im\sigma} = \sum_\alpha (Q^\dagger)^{im}_\alpha \hat{D}^\dagger_{\alpha\sigma}
\tag{7}
$$

where α runs over the 12 molecular orbitals and we are ignoring spin-orbit coupling (SOC) at this stage. Q^α_{im} is the 12×12 transformation matrix from the atomic t_{2g} to the MO spaces. In terms of *global* coordinates (using the same cartesian coordinates for all V sites) it is tabulated in Table I. Note that in actual calculations, since the four V sites are equivalent to each other up to a symmetry operation, Q should be unitarily transformed to a local coordinate system at each V site.

Plugging them into $\hat{H}_K$ yields,

$$
\begin{aligned}
\hat{H}_K = -\sum_{\alpha\beta\gamma\delta} \Big[& (U - 2J) \sum_i \Big\{ \sum_{mm'} (Q^\dagger)^{im}_\alpha (Q^\dagger)^{im'}_\beta Q^\gamma_{im} Q^\delta_{im'} \Big\} \hat{D}^\dagger_{\alpha\uparrow}\hat{D}^\dagger_{\beta\downarrow}\hat{D}_{\gamma\uparrow}\hat{D}_{\delta\downarrow} \\
& + 2J \sum_i \Big\{ \sum_m (Q^\dagger)^{im}_\alpha (Q^\dagger)^{im}_\beta Q^\gamma_{im} Q^\delta_{im} \Big\} \hat{D}^\dagger_{\alpha\uparrow}\hat{D}^\dagger_{\beta\downarrow}\hat{D}_{\gamma\uparrow}\hat{D}_{\delta\downarrow} \\
& + \frac{U - 3J}{2} \sum_i \Big\{ \sum_{m\neq m'} (Q^\dagger)^{im}_\alpha (Q^\dagger)^{im'}_\beta Q^\gamma_{im} Q^\delta_{im'} \Big\} \sum_\sigma \hat{D}^\dagger_{\alpha\sigma}\hat{D}^\dagger_{\beta\sigma}\hat{D}_{\gamma\sigma}\hat{D}_{\delta\sigma} \\
& - J \sum_i \Big\{ \sum_{m\neq m'} (Q^\dagger)^{im}_\alpha (Q^\dagger)^{im'}_\beta Q^\gamma_{im} Q^\delta_{im'} \Big\} \hat{D}^\dagger_{\alpha\uparrow}\hat{D}^\dagger_{\beta\downarrow}\hat{D}_{\gamma\downarrow}\hat{D}_{\delta\uparrow} \\
& - J \sum_i \Big\{ \sum_{m\neq m'} (Q^\dagger)^{im}_\alpha (Q^\dagger)^{im}_\beta Q^\gamma_{im'} Q^\delta_{im'} \Big\} \hat{D}^\dagger_{\alpha\uparrow}\hat{D}^\dagger_{\beta\downarrow}\hat{D}_{\gamma\downarrow}\hat{D}_{\delta\uparrow} \Big].
\end{aligned}
\tag{8}
$$

5

Irreps No. Coeff.

Irreps	No.	V$_1$ (0.4,0.4,0.4)			V$_2$ (0.4,0.6,0.6)			V$_3$ (0.6,0.6,0.4)			V$_4$ (0.6,0.4,0.6)		
		d_{xy}	d_{yz}	d_{xz}	d_{xy}	d_{yz}	d_{xz}	d_{xy}	d_{yz}	d_{xz}	d_{xy}	d_{yz}	d_{xz}
A	1	+1	+1	+1	-1	+1	-1	+1	-1	-1	-1	-1	+1
E	1	+1	+w^1	+w^2	-1	+w^1	-w^2	+1	-w^1	-w^2	-1	-w^1	+w^2
	2	+1	+w^2	+w^1	-1	+w^2	-w^1	+1	-w^2	-w^1	-1	-w^2	+w^1
T^2	1	+1	0	0	+1	0	0	+1	0	0	+1	0	0
	2	0	+1	0	0	+1	0	0	+1	0	0	+1	0
	3	0	0	+1	0	0	+1	0	0	+1	0	0	+1
T_a^1	1	0	+1	-1	0	-1	-1	0	-1	+1	0	+1	+1
	2	+1	0	-1	-1	0	+1	-1	0	-1	+1	0	+1
	3	+1	-1	0	+1	+1	0	-1	-1	0	-1	+1	0
T_b^1	1	0	+1	+1	0	-1	+1	0	-1	-1	0	+1	-1
	2	+1	0	+1	-1	0	-1	-1	0	+1	+1	0	-1
	3	+1	+1	0	+1	-1	0	-1	+1	0	-1	-1	0

TABLE I. Transformation matrix Q_{im}^α from atomic $t_{2\mathrm{g}}$ to molecular orbital basis before normalization, where $w = e^{2\pi i/3}$.

In the above expression, product of Qs can be rewritten as

$$\left(Q^\dagger \otimes Q^\dagger\right)_{\alpha\beta}^{imm'} \equiv (Q^\dagger)_\alpha^{im}(Q^\dagger)_\beta^{im'} \tag{9}$$

$$(Q \otimes Q)_{imm'}^{\gamma\delta} \equiv Q_{im}^\gamma Q_{im'}^\delta, \tag{10}$$

and, since we are considering *local* Coulomb interactions, we are taking direct products of i-subsections ($i=1,\cdots,4$) of Q and $Q^\dagger$ matrices, so that $Q \otimes Q$ (and $Q^\dagger \otimes Q^\dagger$) has dimension of 9×144 for each i when we are considering the full 12-dimensional molecular orbital space.

Since we don't include SOC and the transformation matrices does not have spin indices, all $(Q^\dagger \otimes Q^\dagger) \cdot (Q \otimes Q)$ terms are free of spin components and can be classified into four different kinds; i) $\sum_{mm'} \left(Q^\dagger \otimes Q^\dagger\right)_{\alpha\beta}^{imm'} (Q \otimes Q)_{imm'}^{\gamma\delta}$, ii) $\sum_m \left(Q^\dagger \otimes Q^\dagger\right)_{\alpha\beta}^{imm} (Q \otimes Q)_{imm}^{\gamma\delta}$, iii) $\sum_{m\neq m'} \left(Q^\dagger \otimes Q^\dagger\right)_{\alpha\beta}^{imm'} (Q \otimes Q)_{imm'}^{\gamma\delta}$, and iv) $\sum_{m\neq m'} \left(Q^\dagger \otimes Q^\dagger\right)_{\alpha\beta}^{imm} (Q \otimes Q)_{im'm'}^{\gamma\delta}$. Here case

iii) is just the subtraction of ii) from i).

Computation of the transformation matrix is straightforward, but now all different molecular orbitals can mix even in a simple density-density interaction form (the first three terms in $\hat{H}_K$). However, things become much simpler in the most basic case of considering only the T^2 irrep as the correlated subspace. In that case, all Q_i (and $Q^{\dagger,i}$) become 3×3 identity matrix (with normalization factor 1/2), so that all $Q \otimes Q$ and $Q^\dagger \otimes Q^\dagger$ become 9×9 identity matrix with a prefactor 1/4, so that

$$\text{i)} \sum_{imm'} \left(Q^\dagger \otimes Q^\dagger\right)^{imm'}_{\alpha\beta} (Q \otimes Q)^{\gamma\delta}_{imm'} \to \frac{1}{4}\delta_{\alpha\gamma}\delta_{\beta\delta}, \tag{11}$$

$$\text{ii)} \sum_{im} \left(Q^\dagger \otimes Q^\dagger\right)^{imm}_{\alpha\beta} (Q \otimes Q)^{\gamma\delta}_{imm} \to \frac{1}{4}\delta_{\alpha\gamma}\delta_{\beta\delta}\delta_{\alpha\beta}, \tag{12}$$

$$\text{iii)} \sum_{i,m\neq m'} \left(Q^\dagger \otimes Q^\dagger\right)^{imm'}_{\alpha\beta} (Q \otimes Q)^{\gamma\delta}_{imm'} \to \frac{1}{4}\delta_{\alpha\gamma}\delta_{\beta\delta}(1 - \delta_{\alpha\beta}), \tag{13}$$

$$\text{iv)} \sum_{m\neq m'} \left(Q^\dagger \otimes Q^\dagger\right)^{imm}_{\alpha\beta} (Q \otimes Q)^{\gamma\delta}_{im'm'} \to \frac{1}{4}\delta_{\alpha\beta}\delta_{\gamma\delta}(1 - \delta_{\alpha\gamma}). \tag{14}$$

Hence $\hat{H}_K$, projected onto the MO-T^2 subspace, becomes

$$\hat{H}^{\text{MO}}_K = -\frac{1}{4}\bigg[(U - 2J)\sum_{mm'} \hat{D}^\dagger_{m\uparrow}\hat{D}^\dagger_{m'\downarrow}\hat{D}_{m\uparrow}\hat{D}_{m'\downarrow}$$

$$+2J\sum_m \hat{D}^\dagger_{m\uparrow}\hat{D}^\dagger_{m\downarrow}\hat{D}_{m\uparrow}\hat{D}_{m\downarrow}$$

$$+\frac{U - 3J}{2}\sum_{m\neq m',\sigma} \hat{D}^\dagger_{m\sigma}\hat{D}^\dagger_{m'\sigma}\hat{D}_{m\sigma}\hat{D}_{m'\sigma}$$

$$-J\sum_{m\neq m'} \hat{D}^\dagger_{m\uparrow}\hat{D}^\dagger_{m'\downarrow}\hat{D}_{m\downarrow}\hat{D}_{m'\uparrow}$$

$$-J\sum_{m\neq m'} \hat{D}^\dagger_{m\uparrow}\hat{D}^\dagger_{m\downarrow}\hat{D}_{m'\downarrow}\hat{D}_{m'\uparrow}\bigg]. \tag{15}$$

Note that $\hat{H}^{\text{MO}}_K$ has the exactly same form with the atomic $\hat{H}_K$, except the prefactor 1/4 because of the equidistribution of the MO-T^2 wavefunctions all over the four V sites.

D. Choice of Coulomb interactions and V$_4$ clustering

To check the reliability of employing Ising-like (density-density type) Coulomb interactions in our study, we compare our results presented in the main text with those employing full Coulomb
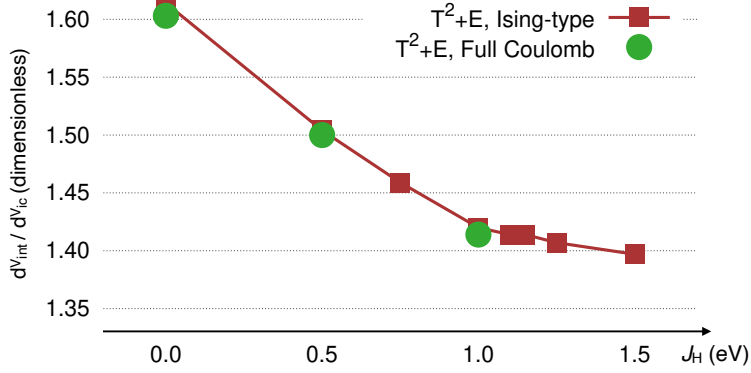
FIG. S1. $d^{\mathrm{V}}_{\mathrm{int}}/d^{\mathrm{V}}_{\mathrm{ic}}$ from different choices (Ising-like and full Coulomb) of on-site Coulomb interactions as a function of J_{H}, where MO-$\{T^2 \oplus E\}$ is employed as the correlated orbitals. Note that two results show similar $d^{\mathrm{V}}_{\mathrm{int}}/d^{\mathrm{V}}_{\mathrm{ic}}$ values and same tendency with respect to the increasing J_{H}.

interactions. For the comparison we chose the MO-$\{T^2 \oplus E\}$ as our correlated subspace because computational costs using full Coulomb interactions in the MO-$\{T^2 \oplus E \oplus T_a^1\}$ configuration exceeds our hardware limit. We would like to argue that, if the tendency of $d^{\mathrm{V}}_{\mathrm{int}}/d^{\mathrm{V}}_{\mathrm{ic}}$ as a function of J_{H} is consistent across both choices of Coulomb interactions in the MO-$\{T^2 \oplus E\}$ configuration, then it should be so in the MO-$\{T^2 \oplus E \oplus T_a^1\}$ setup as well. This is because effects of Hund's coupling are most dominant within the MO-$\{T^2 \oplus E\}$ subspace, and although the inclusion of the T_a^1 orbital is crucial in obtaining realistic value of $d^{\mathrm{V}}_{\mathrm{int}}/d^{\mathrm{V}}_{\mathrm{ic}}$, electron occupation in the T_a^1 orbital remains small (< 0.1) even in the case of $J_{\mathrm{H}} > 1$ eV.

Figure S1 shows the comparison of calculated $d^{\mathrm{V}}_{\mathrm{int}}/d^{\mathrm{V}}_{\mathrm{ic}}$ between two choices of Coulomb interactions: Ising-like and full. It can be seen that the choice of Coulomb interactions does not make any qualitative differences. While the choice of Ising-like interactions breaks rotational symmetry in the magnetic sector and may affect magnetic properties and metal-insulator transition behaviors, its effects on structural degrees of freedom in our case seems less significant.

E. On-site and inter-site self-energies

In this section the role of the Hund's coupling is discussed in terms of the real space representation of the self-energy. Here we focus on the $T^2 \oplus E$ subspaces and their self-energies. Similar analysis can be done with other MO subspaces, however, for the purpose of discussing the role of J_{H} it seems that $T^2 \oplus E$ should suffice.

8

In our calculations the cluster self-energies are diagonalized within the MO representation. When back-transformed into the atomic orbital basis representation, on-site (local) and inter-site (non-local) self-energies within the V_4 tetrahedron can be obtained. In the simplest case with the correlated MO-T^2 triplet only, the form of the self-energy in the atomic representation becomes simple; Namely, in the four-site real-space representation (four sites $\otimes$ atomic t_{2g}), all the on-site and inter-site blocks are enforced to be identical due to the choice of the T^2 correlated orbitals when the cubic and time-reversal symmetries are present, so that

$$\mathbf{\Sigma}\left[T^2\right](\omega) \equiv \frac{1}{4}\hat{\Sigma}^{T^2}(\omega)\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}, \tag{16}$$

where each 3×3 block $\hat{\Sigma}^{T^2} = \Sigma^{T^2} \times \hat{I}_{3\times 3}$ in the atomic t_{2g} space (d_{xy}, d_{yz}, and d_{xz}), $\hat{I}_{3\times 3}$ is an identity matrix of dimension 3, and the frequency ω can be either real or imaginary. Note that Σ^{T^2} is the diagonal self-energy in the T^2-MO representation, and that the prefactor $\frac{1}{4}$ in Eq. (16) is the one appearing in Eq. (15). Here we choose the same global coordinate in defining the t_{2g} orbitals at all sites, and proper coordinate transforms should be applied to each block when represented in local coordinates ($\hat{\Sigma}^{T^2}_{ij} \to (Q^\dagger)_{gi}\hat{\Sigma}^{T^2}_{ij}Q_{jg}$, where the transformation Q_{ig} is made from the global to the site-i local coordinates). Fig. S2 plots the real and imaginary parts of Σ^{T^2} in the real frequency space, showing a pole in the imaginary part inside the Mott gap.

From this real-space representation of the self-energy, the implication of choosing only the T^2 MO as correlated orbitals becomes clearer; i) it introduces the inter-site self-energy in addition to the on-site counterpart, and ii) it prevents the correlations from becoming more local by enforcing the on-site and inter-site self-energies to be identical. The latter, especially, can be a serious issue when the size of the correlations that favor the formation of the local moments, *e.g.* the Hund's coupling, becomes comparable to that of inter-site hopping.

Next, the form of self-energy in the $T^2 \oplus E$ is as presented below:

$$\mathbf{\Sigma}\left[T^2 \oplus E\right](\omega) = \mathbf{\Sigma}\left[T^2\right](\omega) + \mathbf{\Sigma}\left[E\right](\omega), \tag{17}$$
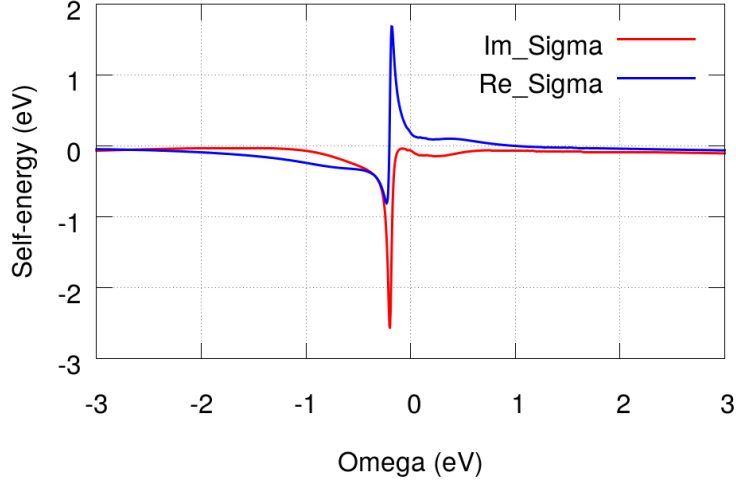
FIG. S2. Real (blue) and imaginary (red) part of $\Sigma^{T^2}(\omega)$ after the analytic continuation from the imaginary frequency space.

where the T^2-part of the self-energy is shown in Eq. (16). $\boldsymbol{\Sigma}\left[E\right](\omega)$ is as follows;

$$\boldsymbol{\Sigma}\left[E\right](\omega) \equiv \begin{pmatrix} \hat{\Sigma}_{11} & \hat{\Sigma}_{12} & \hat{\Sigma}_{13} & \hat{\Sigma}_{14} \\ \hat{\Sigma}_{12}^T & \hat{\Sigma}_{22} & \hat{\Sigma}_{23} & \hat{\Sigma}_{24} \\ \hat{\Sigma}_{13}^T & \hat{\Sigma}_{23}^T & \hat{\Sigma}_{33} & \hat{\Sigma}_{34} \\ \hat{\Sigma}_{14}^T & \hat{\Sigma}_{24}^T & \hat{\Sigma}_{34}^T & \hat{\Sigma}_{44} \end{pmatrix}. \tag{18}$$

Here the on-site parts $\hat{\Sigma}_{ii}$ are

$$\hat{\Sigma}_{ii} \equiv \Sigma^E(\omega)\left(\frac{1}{6}\hat{I}_{3\times 3} + \frac{1}{12}\hat{\Delta}_{ii}\right), \tag{19}$$

where $\Sigma^E(\omega)$ is the self-energy for the E doublet in the MO representation, and $\hat{\Delta}_{ii}$ determines the direction of the "*trigonal crystal fields*" to $t_{2\mathrm{g}}$ orbitals at each V site, exerted by $\frac{1}{12}\Sigma^E(\omega)\hat{\Delta}_{ii}$. Namely, if the VS_6 octahedron surrounding site 1 is trigonally distorted along the cubic [111] direction with respect to the global Cartesian coordinate (*i.e.*, if the site 1 and the center of the V_4 cluster are on the same [111] line), then

$$\hat{\Delta}_{11} = \begin{pmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{pmatrix}. \tag{20}$$
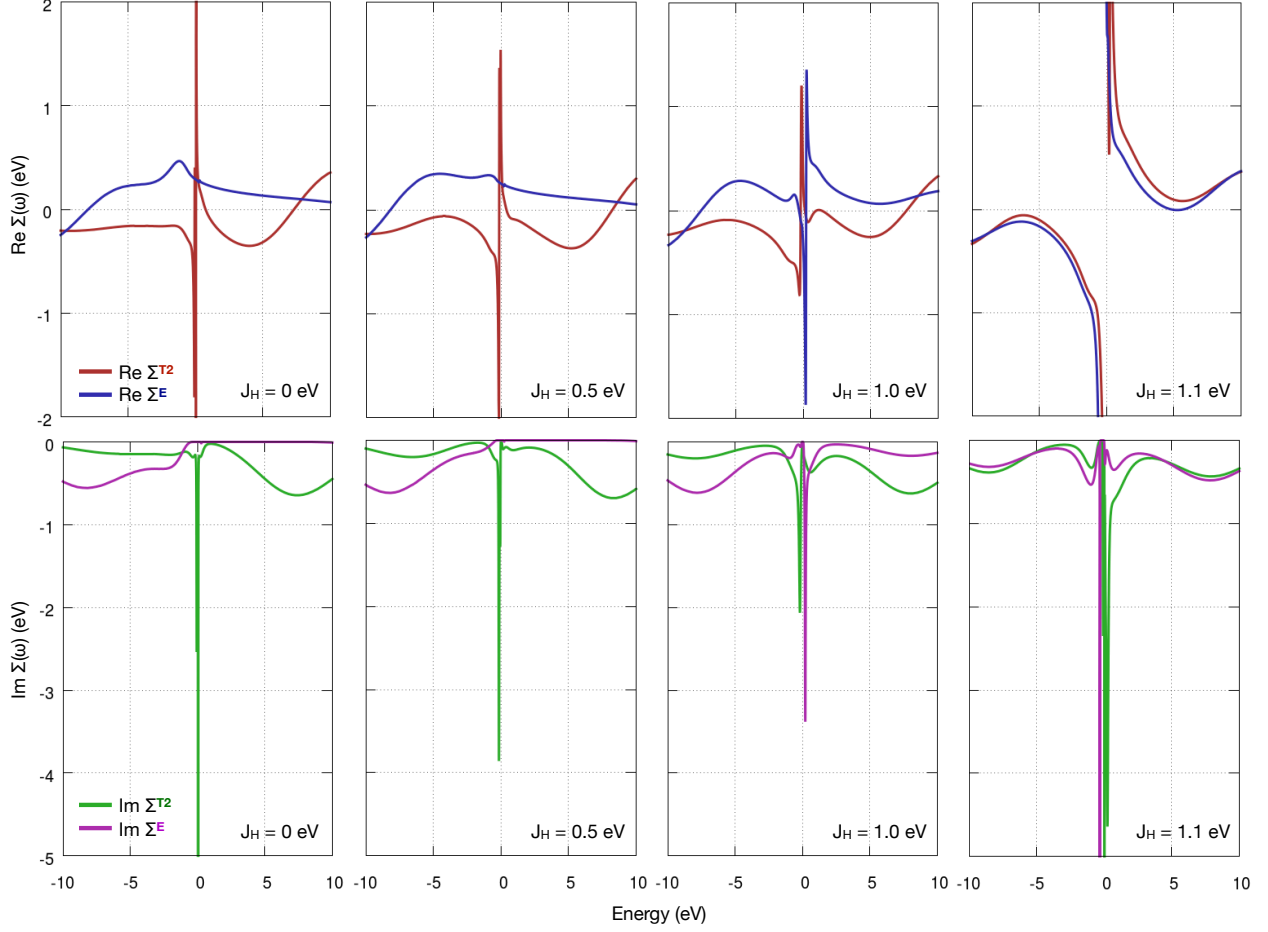
10

FIG. S3. Real and imaginary part of $\Sigma^{T^2,E}(\omega)$ after the analytic continuation from the imaginary frequency space. Top and bottom panels depict real and imaginary parts, respectively. From left to right, size of the Hund's coupling J_H is enhanced ($J_H = 0, 0.5, 1.0, 1.1$ eV). Note that the high-spin configuration is stabilized at $J_H = 1.1$ eV.

Other $\hat{\Delta}_{ii}$, for a coordinate choice, should be as follows,

$$\hat{\Delta}_{22} = \begin{pmatrix} 0 & +1 & -1 \\ +1 & 0 & +1 \\ -1 & +1 & 0 \end{pmatrix}, \hat{\Delta}_{33} = \begin{pmatrix} 0 & +1 & +1 \\ +1 & 0 & -1 \\ +1 & -1 & 0 \end{pmatrix}, \hat{\Delta}_{44} = \begin{pmatrix} 0 & -1 & +1 \\ -1 & 0 & +1 \\ +1 & +1 & 0 \end{pmatrix}. \qquad (21)$$

Note that this is the coordinate choice that was adopted in this work.

11

The inter-site component $\hat{\Sigma}_{ij}$ has a similar form; $\hat{\Sigma}_{ij} \equiv \frac{1}{12}\Sigma^E(\omega)\hat{O}_{ij}$, where

$$
\hat{O}_{12} = \begin{pmatrix} -2 & -1 & +1 \\ +1 & +2 & +1 \\ +1 & -1 & -2 \end{pmatrix}, \hat{O}_{13} = \begin{pmatrix} +2 & +1 & +1 \\ -1 & -2 & +1 \\ -1 & +1 & -2 \end{pmatrix}, \hat{O}_{23} = \begin{pmatrix} -2 & +1 & -1 \\ +1 & -2 & -1 \\ +1 & +1 & +2 \end{pmatrix},
$$

$$
\hat{O}_{23} = \begin{pmatrix} -2 & -1 & -1 \\ -1 & -2 & +1 \\ +1 & -1 & +2 \end{pmatrix}, \hat{O}_{24} = \begin{pmatrix} +2 & -1 & +1 \\ +1 & -2 & -1 \\ -1 & -1 & -2 \end{pmatrix}, \hat{O}_{34} = \begin{pmatrix} -2 & +1 & -1 \\ -1 & +2 & +1 \\ -1 & -1 & -2 \end{pmatrix}. \tag{22}
$$

Combining (16-22), the site-orbital resolved self-energies in the $T^2 \oplus E$ case is as follows.

i) *On-site (diagonal blocks), between same orbitals*: $\left[\frac{1}{4}\Sigma^{T^2}(\omega) + \frac{1}{6}\Sigma^E(\omega) \right] \hat{I}_{3\times3}$,

ii) *On-site (diagonal blocks), between different orbitals*: $\frac{1}{12}\Sigma^E(\omega)\hat{\Delta}_{ii}$,

iii) *Inter-site ($i \neq j$ blocks)*: $\frac{1}{4}\Sigma^{T^2}(\omega)\hat{I}_{3\times3} + \frac{1}{12}\Sigma^E(\omega)\hat{O}_{ij}$.

Here, we note in passing that $\frac{1}{12}\Sigma^E$ is small compared to other terms when J_{H} is not large (< 1 eV), so that terms i) and iii) are dominant contributions, and that the balance between the terms i) and iii) determines whether it is locally (on-site) or non-locally (inter-site) correlated. Plugging (22) into the case iii) above yields an explicit expression of the ij-block of $\mathbf{\Sigma}\left[T^2 \oplus E\right]$. For example, the block between the site 1 and 2 is as follows,

$$
\mathbf{\Sigma}\left[T^2 \oplus E\right]_{12} = \begin{pmatrix} \frac{1}{4}\Sigma^{T^2} - \frac{1}{6}\Sigma^E & -\frac{1}{12}\Sigma^E & +\frac{1}{12}\Sigma^E \\ +\frac{1}{12}\Sigma^E & \frac{1}{4}\Sigma^{T^2} + \frac{1}{6}\Sigma^E & +\frac{1}{12}\Sigma^E \\ +\frac{1}{12}\Sigma^E & -\frac{1}{12}\Sigma^E & \frac{1}{4}\Sigma^{T^2} - \frac{1}{6}\Sigma^E \end{pmatrix}, \tag{23}
$$

where the plus and minus signs in the diagonal components are colored in blue and red to emphasize terms where Σ^{T^2} and Σ^E are adding up and cancelling out, respectively. Among the three diagonal components, the central term ($\frac{1}{4}\Sigma^{T^2} + \frac{1}{6}\Sigma^E$) is between the d_{yz} orbitals at V site 1 and 2, which are forming a strong σ-type direct overlap, while the other two $\frac{1}{4}\Sigma^{T^2} - \frac{1}{6}\Sigma^E$ are contributing to the δ-like weak overlap between the $d_{xy,xz}$ orbitals. Interestingly, the inclusion of Σ^E (and J_{H}) affects the inter-site self-energies in an opposite way depending on the orbitals; while the imaginary part of $\frac{1}{4}\Sigma^{T^2} + \frac{1}{6}\Sigma^E$ is enhanced by the nonzero Σ^E (because causal self-energies should always have negative imaginary parts), it is canceled out in $\frac{1}{4}\Sigma^{T^2} - \frac{1}{6}\Sigma^E$. This implies that the presence of Σ^E selectively enhances the singlet moment formation within the stronger σ-bonding,

while reducing inter-site correlations in other bondings. In addition, depending on the sign of the real parts of Σ^{T^2} and Σ^E, one can either enhance or suppress the real part of the self-energy.

Fig. S3 show the evolution of $\Sigma^{T^2,E}(\omega)$ as a function of the Hund's coupling J_H. Note that the relative signs of the real part of $\Sigma^{T^2,E}(\omega)$ tend to be opposite when J_H is small, but increasing J_H drives them to be the same. Just after the crossover to the high-spin state happens ($J_\mathrm{H} = 1.1$ eV), both the $\mathrm{Re}\Sigma^{T^2,E}(\omega)$ show very similar behavior. This is because of the development of the pole in Σ^E, signaling the formation of the E local moments, as shown in the lower panels of Fig. S3. As the system goes into the high-spin configuration, both the $\mathrm{Im}\Sigma^{T^2,E}$ should similarly show a well-defined pole, then the shapes of $\mathrm{Re}\Sigma^{T^2,E}(\omega)$ should become similar to each other because of the Kramers-Kronig relation. Hence $\frac{1}{4}\Sigma^{T^2} - \frac{1}{6}\Sigma^E$ within $\mathbf{\Sigma}\left[T^2 \oplus E\right]_{ij}$ tends to cancel better as J_H becomes larger. Since the diagonal parts of the inter-site self energies are most dominant contributions, and we have two $\frac{1}{4}\Sigma^{T^2} - \frac{1}{6}\Sigma^E$ terms compared to just one $\frac{1}{4}\Sigma^{T^2} + \frac{1}{6}\Sigma^E$, the overall self-energy correction to the inter-site hopping terms becomes weaker as the Hund's coupling becomes enhanced. This is consistent with the observation in the main text that increasing J_H suppresses the degree of V_4 clustering, and that while U enhanced the inter-site correlation via Σ^{T^2}, J_H reduces it by introducing Σ^E that cancels Σ^{T^2} out.

[1] G. Kresse and J. Hafner, "*Ab initio* molecular dynamics for liquid metals," Phys. Rev. B **47**, 558–561 (1993).

[2] G. Kresse and J. Furthmüller, "Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set," Phys. Rev. B **54**, 11169–11186 (1996).

[3] D. M. Ceperley and B. J. Alder, "Ground state of the electron gas by a stochastic method," Phys. Rev. Lett. **45**, 566–569 (1980).

[4] John P. Perdew, Kieron Burke, and Matthias Ernzerhof, "Generalized gradient approximation made simple," Phys. Rev. Lett. **77**, 3865–3868 (1996).

[5] Gábor I. Csonka, John P. Perdew, Adrienn Ruzsinszky, Pier H. T. Philipsen, Sébastien Lebègue, Joachim Paier, Oleg A. Vydrov, and János G. Ángyán, "Assessing the performance of recent density functionals for bulk solids," Phys. Rev. B **79**, 155107 (2009).

[6] Jianwei Sun, Adrienn Ruzsinszky, and John P. Perdew, "Strongly constrained and appropriately normed semilocal density functional," Phys. Rev. Lett. **115**, 036402 (2015).

[7] S. L. Dudarev, G. A. Botton, S. Y. Savrasov, C. J. Humphreys, and A. P. Sutton, "Electron-energy-loss spectra and the structural stability of nickel oxide: An LSDA+U study," Phys. Rev. B **57**, 1505–1509 (1998).

[8] Jochen Heyd, Gustavo E. Scuseria, and Matthias Ernzerhof, "Hybrid functionals based on a screened coulomb potential," J. Chem. Phys. **118**, 8207–8215 (2003).

[9] Jochen Heyd, Gustavo E. Scuseria, and Matthias Ernzerhof, "Erratum: Hybrid functionals based on a screened Coulomb potential [J. Chem. Phys.118, 8207 (2003)]," The Journal of Chemical Physics **124**, 219906 (2006), http://dx.doi.org/10.1063/1.2204597.

[10] Kristjan Haule, Chuck-Hou Yee, and Kyoo Kim, "Dynamical mean-field theory within the full-potential methods: Electronic structure of CeIrIn$_5$, CeCoIn$_5$, and CeRhIn$_5$," Phys. Rev. B **81**, 195107 (2010).

[11] P. Blaha, K. Schwarz, G. K. H. Madsen, D. Kvasnicka, and J. Luitz, *WIEN2k, An Augmented Plane Wave + Local Orbitals Program for Calculating Crystal Properties* (Karlheinz Schwarz, Techn. Universität Wien, Austria, 2001).

[12] Kristjan Haule and Gheorghe L. Pascut, "Forces for structural optimizations in correlated materials within a dft+embedded dmft functional approach," Phys. Rev. B **94**, 195146 (2016).

[13] Kristjan Haule and Turan Birol, "Free energy from stationary implementation of the DFT + DMFT functional," Phys. Rev. Lett. **115**, 256402 (2015).

[14] Regina Pocha, Dirk Johrendt, and Rainer Pttgen, "Electronic and Structural Instabilities in GaV$_4$S$_8$ and GaMo$_4$S$_8$," Chemistry of Materials **12**, 2882–2887 (2000).

[15] Kristjan Haule, "Quantum monte carlo impurity solver for cluster dynamical mean-field theory and electronic structure calculations with adjustable cluster base," Phys. Rev. B **75**, 155113 (2007).

[16] Kristjan Haule, "Exact double counting in combining the dynamical mean field theory and the density functional theory," Phys. Rev. Lett. **115**, 196403 (2015).

[17] Kristjan Haule, Turan Birol, and Gabriel Kotliar, "Covalency in transition-metal oxides within all-electron dynamical mean-field theory," Phys. Rev. B **90**, 075136 (2014).