Supplementary Material for "Hamiltonian parameter inference from resonant inelastic x-ray scattering with active learning"

Marton K. Lajer,[1,*] Xin Dai,[2] Kipton Barros,[3]
Matthew R. Carbone,[2] S. Johnston,[4,5,†] and M. P. M. Dean[1,4,‡]

[1] *Condensed Matter Physics and Materials Science Department, Brookhaven National Laboratory, Upton, New York 11973, USA*
[2] *Computing and Data Sciences Directorate, Brookhaven National Laboratory, Upton, New York 11973, USA*
[3] *Theoretical Division and CNLS, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA*
[4] *Department of Physics and Astronomy, The University of Tennessee, Knoxville, Tennessee 37996, USA*
[5] *Institute for Advanced Materials and Manufacturing, University of Tennessee, Knoxville, Tennessee 37996, USA*

(Dated: October 20, 2025)

S1. DETERMINING SYMMETRY LABELS

A. Quantum numbers

Let $\mathcal{H}$ be a Hilbert space, and Q_i, $i \in \{1, 2, \dots n_q\}$ mutually commuting operators acting in $\mathcal{H}$ such that

$$[Q_i, Q_j] = 0, \quad \forall i, j.$$

The set $\{Q_i\}$ has a common set of eigenvectors $|\{q_i\}; r_i\rangle$ labeled by the set of eigenvalues $Q_j |\{q_i\}; r_i\rangle = q_j |\{q_i\}; r_i\rangle$. The numbers q_i are called quantum numbers. The index r_i runs through the degenerate subspaces with completely coinciding quantum numbers $\{q_i\}$.

EDRIXS works in the Fock basis where basis states are labeled by the occupation numbers of individual orbitals. However, we are often interested in the eigenbasis of operators like S_z, S^2, L_z, L^2, $J_z = (L_z + S_z)$ or J^2. These operators don't all commute with the occupation number operators, so a basis vector in the occupation basis will not generally have a definite angular momentum eigenvalue.

Let us construct an auxiliary operator A such that

$$A = \sum_{i=1}^{n_q} a_i Q_i$$

with a_i chosen to be algebraically independent. For example, a convenient choice is to take $a_i = \eta\sqrt{p_i}$, where η is an arbitrary constant and p_i is the sequence such that $p_1 = 1$ and p_i is the $i-1$ smallest prime number for $i > 1$. If the matrix elements of Q_i are computed in the occupation basis, the unitary matrix built from the eigenvectors of A provides the basis transform to the basis with definite symmetry labels. The set of q values labeling a degenerate subspace can by obtained from the eigenvalues of A:

$$A \left|\phi_i^A\right\rangle = \eta \sum_{i=1}^{n_q} \sqrt{p_i} q_i \left|\phi_i^A\right\rangle$$

using e.g. simple table lookup. Alternatively, one can act the symmetry operators directly, $Q_i \left|\phi_j^A\right\rangle = q_i^{(j)} \left|\phi_j^A\right\rangle$.

B. Discrete symmetry labels

To decompose the degenerate subspaces of A further, we consider random symmetric matrices $R_{\{q_i\}}$ with matrix elements sampled from an uniform distribution. We intend to build matrices invariant with respect to a discrete symmetry group G. To this end, we construct the operator

$$\bar{R}_{\{q_i\}} = \frac{1}{|G|} \sum_{g \in G} D(g) R_{\{q_i\}} D^\dagger(g)$$

* mlajer@bnl.gov
† sjohn145@utk.edu
‡ mdean@bnl.gov

It is easy to see that $\bar{R}_{\{q_i\}}$ is invariant with respect to G by conjugating it with an arbitrary group element $g' \in G$:

$$\begin{aligned} D(g')\,\bar{R}_{\{q_i\}}D^\dagger(g') &= \frac{1}{|G|}\sum_{g\in G} D(g')\,D(g)\,R_{\{q_i\}}D^\dagger(g)\,D^\dagger(g') \\ &= \frac{1}{|G|}\sum_{g\in G} D(g'g)\,R_{\{q_i\}}D^\dagger(g'g) \\ &= \frac{1}{|G|}\sum_{h\in G} D(h)\,R_{\{q_i\}}D^\dagger(h) \\ &= \bar{R}_{\{q_i\}}. \end{aligned}$$

Since $\bar{R}_{\{q_i\}}$ commutes with all group elements, its eigenvectors are restricted to transform within its degenerate subspaces. A priori it could happen that one such subspace contains multiple irreducible representations (irreps). However, as we constructed R with random matrix elements, it is guaranteed (with probability ~ 1 up to machine precision) that $\bar{R}_{\{q_i\}}$ does not have accidental degeneracies in its eigenvalues apart from that strictly required by its group invariance. The unlikely event of an accidental degeneracy is under control (see below) and can be corrected by repeating with another random matrix R. Therefore the eigenvectors of $\bar{R}_{\{q_i\}}$ can be assumed to transform according to individual irreps of G. In turn, the eigenvectors can be labeled by the particular irrep copy they belong to.

Given the character table of the group G, we construct the square matrix T with matrix elements corresponding to the entries of the character table. We then pick an exemplar group element from each conjugate class and calculate the trace of their representations in the degenerate subspace of $\bar{R}_{\{q_i\}}$ with eigenvalue λ. Thus, in each subspace λ, we get a vector $\mathcal{X}_\lambda$ of characters. The multiplicity of each irrep in the subspace is given by the expression

$$\mu_\lambda = \mathcal{X}_\lambda T^{-1}. \tag{1}$$

If the degenerate subspace indeed corresponds to an irrep, the vector μ is a standard basis vector with one element being 1 and the rest are zeros. We can read the type of the irrep by comparing the position of the one in μ and the arrangement of the rows of the character table.

Let us consider a Hamiltonian H acting in the Hilbert space $\mathcal{H}$. We insist that H is not necessarily invariant under the symmetry operations generated by Q_i:

$$[H, Q_i] \neq 0.$$

We also do not require invariance with respect to the finite group G. Instead, after we obtain the eigenvectors of $H|\psi\rangle = E_\psi|\psi\rangle$, we calculate the squares of overlaps between $|\psi\rangle$ and the completely annotated basis vectors $|\phi_{q_i;\lambda;r}\rangle$

$$w_{\psi,q_i;\lambda} = \sum_{r=1}^{\dim D_\lambda} |\langle\psi|\phi_{q_i;\lambda;r}\rangle|^2 . \tag{2}$$

Since the wavefunction is normalized and the annotated vectors span the Hilbert space $\mathcal{H}$,

$$\sum_{q_i;\lambda} w_{\psi,q_i;\lambda} = 1. \tag{3}$$

The weight of any combination of symmetry labels in $|\psi\rangle$ is easily calculated by summing up the appropriate subset of elementary weights $w_{\psi,q_i;\lambda}$.

We note that projections of $|\psi\rangle$ to isotypic subspaces of the group G corresponding to an irrep Γ could also be calculated by the projection formula (see e.g. Ref. [1])

$$\sum_{\substack{\lambda\in\text{copies}\\ \text{of irrep }\Gamma}} w_{\psi,q_i;\lambda} = \left| \frac{\dim\Gamma}{|G|}\sum_{g\in G} \chi^\Gamma(g)^*\,\langle\psi|D_{q_i}(g)|\psi\rangle \right|^2 , \tag{4}$$

where $D_{q_i}(g)$ is the (generally reducible) representation of G that acts in the degenerate subspace $\{q_i\}$ of the matrix A. This approach avoids the construction of the commutant $\bar{R}$, but does not give access to weights in individual irrep copies. Such extra information can sometimes be useful, as we discuss in Section S1 D.

This concludes our general strategy of eigenstate annotation. We stress that the commutant method is applicable for any finite group.[1]

[1] Although tangential to the present discussion, we remark that this method entails a didactic, systematic way to obtain character

C. The octahedral group

We focus on the octahedral group Oh, and specifically its orientation-preserving subgroup O. The group Oh is of order 48, while the group O has 24 elements. The reason of our focus is that we intend to annotate the states of the initial Hamiltonian, which are linear combinations of d-shell excitations. Therefore, all eigenfunctions are invariant under inversion.

The group elements of the octahedral group take the form

$$g_\alpha = e^{\mathrm{i}\alpha_x L_x + \mathrm{i}\alpha_y L_y + \mathrm{i}\alpha_z L_z}.$$

Elements of the group O are organized into 5 conjugacy classes as follows.

1. ($6C_4$ – order 4) 6 rotations about the cube edges with angles $n\frac{\pi}{2}$, $n \in \{1,3\}$:

$$\alpha_{3n-2} = n\frac{\pi}{2}(1,0,0), \quad \alpha_{3n-1} = n\frac{\pi}{2}(0,1,0), \quad \alpha_{3n} = n\frac{\pi}{2}(0,0,1).$$

2. ($3C_2 \sim (C_4)^2$ – order 2) 3 rotations about the cube edges with angle π:

$$\alpha_4 = \pi(1,0,0), \quad \alpha_5 = \pi(0,1,0), \quad \alpha_6 = \pi(0,0,1).$$

3. ($8C_2$ – order 2) 8 rotations about the face diagonals with angle π:

$$\begin{aligned}\alpha_{10} &= \pi(1,1,0), \quad \alpha_{11} = \pi(1,-1,0), \quad \alpha_{12} = \pi(1,0,1)\\ \alpha_{13} &= \pi(1,0,-1), \quad \alpha_{14} = \pi(0,1,1), \quad \alpha_{15} = \pi(0,1,-1).\end{aligned}$$

4. ($6C_3$ – order 3) 6 rotations about the body diagonals with angle $\frac{2\pi}{3}n$, $n \in \{1,2\}$:

$$\begin{aligned}\alpha_{4n+12} &= \frac{2\pi}{3}n(1,1,1), \quad \alpha_{4n+13} = \frac{2\pi}{3}n(1,1,-1),\\ \alpha_{4n+14} &= \frac{2\pi}{3}n(1,-1,1), \quad \alpha_{4n+15} = \frac{2\pi}{3}n(-1,1,1).\end{aligned}$$

5. (E – order 1) the identity,

$$\alpha_0 = (0,0,0).$$

The remaining 24 elements and further 5 conjugacy classes of Oh are constructed multiplying the vectors above by -1.

Since there are 5 conjugacy classes in the group O, there are 5 inequivalent irreps. The character table of the group O is shown on Table S1.

The inverse matrix T^{-1} of Eq. (1) takes the form

$$T^{-1} = \frac{1}{24}\begin{pmatrix} 1 & 8 & 6 & 6 & 3\\ 1 & 8 & -6 & -6 & 3\\ 2 & -8 & 0 & 0 & 6\\ 3 & 0 & -6 & 6 & -3\\ 3 & 0 & 6 & -6 & -3 \end{pmatrix}.$$

TABLE S1. Character table of the group O $\cong \mathcal{S}_4$

	E	$8C_3$	$6C_2$	$6C_4$	$3C_2$
A_{1g}	1	1	1	1	1
A_{2g}	1	1	-1	-1	1
E_g	2	-1	0	0	2
T_{1g}	3	0	-1	1	-1
T_{2g}	3	0	1	-1	-1

tables from multiplication tables of finite matrices. One then starts by constructing the regular representation of dimension $|G|$. All inequivalent irreps are present in the regular representation. Solving for the spectrum of R' provides the invariant subspaces of the group. The characters are then obtained from the traces of group elements projected into the invariant subspaces. This method allows the computation of character tables for groups of order up to a few tens of thousands on a PC, depending on available memory.

D. Weights in individual irreducible representation copies

Let G be a finite group with irreps labeled by α. Let D be a (reducible) representation of the group. We can decompose D into isotypic blocks $D_\alpha^{m_\alpha}$ as

$$D = \bigoplus_\alpha D_\alpha^{m_\alpha},$$

where D_α denotes an irrep and m_α is its multiplicity in D, that is

$$D_\alpha^{m_\alpha} = D_\alpha^{(1)} \oplus D_\alpha^{(2)} \oplus \cdots \oplus D_\alpha^{(m_\alpha)}.$$

The decomposition of D into isotypic blocks is unique. However, the decomposition of an isotypic block into its component irreps is only fixed up to $O(m_\alpha)$ orthogonal transformations. To see this, let us consider an arbitrary group element $g \in G$. The representation of g according to the irrep D_α is $D_\alpha(g)$. If the irrep is n_α dimensional, then $D_\alpha(g)$ is an $n_\alpha \times n_\alpha$matrix. In other words, we can introduce a set of n_α orthonormal vectors $e_{\alpha,i}$, $i \in \{1, 2, \dots n_\alpha\}$, so that

$$D_\alpha(g) \equiv \sum_{ab} (e_{\alpha,a} \circ e_{\alpha,b}) [D_\alpha(g)]_{ab}.$$

A vector v_α is said to transform according to the irreducible representation D_α, if

$$v_\alpha = \sum_{i=1}^{\dim D_\alpha} e_{\alpha,i} v_{\alpha,i}.$$

The transformation of v by the group element g is then written as

$$D_\alpha(g)\, v = \sum_{ab=1}^{\dim D_\alpha} (e_{\alpha,a} \circ e_{\alpha,b}) [D_\alpha(g)]_{ab} \sum_{i=1}^{\dim D_\alpha} e_{\alpha,i} v_i = \sum_{a,i=1}^{\dim D_\alpha} e_{\alpha,a} [D_\alpha(g)]_{ai} v_i. \tag{5}$$

Likewise, the representation of g in D is written as

$$D(g) = \sum_\alpha \sum_{n=1}^{m_\alpha} \sum_{ab=1}^{\dim D_\alpha} \left(e_{\alpha,a}^{(n)} \circ e_{\alpha,b}^{(n)}\right) [D_\alpha(g)]_{ab}.$$

The core of the redundancy is that within each isotypic block, we can introduce a new set of basis vectors,

$$\tilde{e}_{\alpha,a}^{(k)} = \sum_{i=1}^{m_\alpha} c_i^{(k)} e_{\alpha,a}^{(i)}; \quad \sum_{i=1}^{m_\alpha} \left|c_i^{(k)}\right|^2 = 1.$$

The normalization condition of the coefficients $c_i^{(k)}$ ensures that the set of basis vectors $\left\{\tilde{e}_{\alpha,a}^{(k)}\right\}$ is also orthonormal. Let us consider a generic vector

$$v_{m\alpha} = \sum_{n=1}^{m_\alpha} \sum_{i=1}^{\dim D_\alpha} e_{\alpha,i}^{(n)} v_{m\alpha,i}^{(n)}.$$

The projection amplitude of the vector $v_{m\alpha}$ onto the irrep $D_\alpha^{(k)}$ is

$$w_k(v_{m\alpha}) = \sum_{j=1}^{\dim D_n} e_{\alpha,j}^{(k)} \sum_{n=1}^{m_\alpha} \sum_{i=1}^{\dim D_\alpha} e_{\alpha,i}^{(n)} v_{m\alpha,i}^{(n)} = \sum_{j=1}^{\dim D_\alpha} v_{m\alpha,j}^{(k)}. \tag{6}$$

What is the largest amplitude of $v_{m\alpha}$ in a single irrep? To answer this, we need to take into account the redundancy of the definition of irreps. The amplitude on a transformed irrep $\tilde{D}_\alpha^{(k)}$takes the form

$$\tilde{w}_k(v_{m\alpha}) = \sum_{j=1}^{\dim D_n} \tilde{e}_{\alpha,j}^{(k)} \sum_{n=1}^{m_\alpha} \sum_{i=1}^{\dim D_\alpha} e_{\alpha,i}^{(n)} v_{m\alpha,i}^{(n)} = \sum_{l=1}^{m_\alpha} c_l^{(k)} \sum_{j=1}^{\dim D_n} v_{m\alpha,j}^{(l)} = \sum_{l=1}^{m_\alpha} c_l^{(k)} w_l(v_{m\alpha}). \tag{7}$$

We would like to calculate

$$\max_{\|c\|^2=1} \sum_{l=1}^{m_\alpha} c_l w_l\left(v_{m\alpha}\right) = \max_{\|c\|^2=1} \sum_{j=1}^{\dim D_\alpha} \sum_{l=1}^{m_\alpha} v^{(l)}_{m_\alpha,j} c_l = \left(\max_{\|c\|^2=1} \left(\sum_{j=1}^{\dim D_\alpha} \sum_{l=1}^{m_\alpha} v^{(l)}_{m\alpha,j} c_l \right)^2 \right)^{\frac{1}{2}} . \tag{8}$$

Introducing $V_{lj} = v^{(l)}_{m\alpha,j}$, we then obtain

$$\max_{\|c\|^2=1} \sum_{l=1}^{m_\alpha} c_l w_l\left(v_{m\alpha}\right) = \left[\max_{\|c\|^2=1} \sum_{l,p=1}^{m_\alpha} c_l \left(\sum_{i,j=1}^{\dim D_\alpha} V_{lj} V^\dagger_{ip} \right) c_p \right]^{\frac{1}{2}} = \left[\lambda_{\max} \left(\sum_{i,j=1}^{\dim D_\alpha} V_{lj} V^\dagger_{ip} \right) \right]^{\frac{1}{2}} . \tag{9}$$

This result is independent of the initial basis chosen and well-defined.

It can be useful to know the maximal weight of a single irrep copy in an eigenvector. For example, this can provide useful information on the shape of the electron density. If the (multiparticle) eigenvector is dominated by a single irrep, the electron density will be reminiscent of the geometry of a single-particle wavefunction.

S2. FURTHER DETAILS OF THE FITTING PROCEDURE AND RESULTS

In this section, we provide several additional plots that characterize the performance and details of our fitting procedure.

A. $NiCl_2$

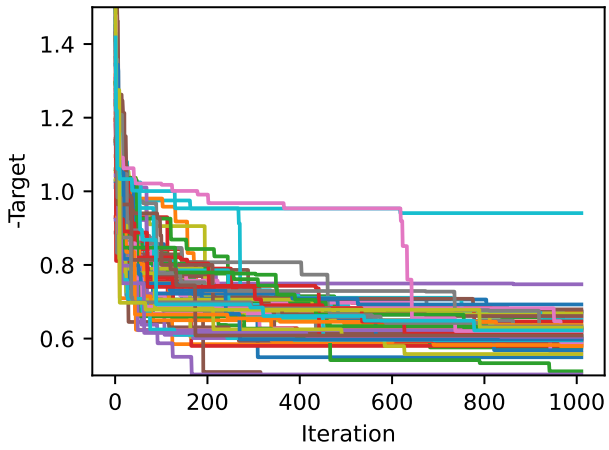

FIG. S1. $NiCl_2$: Decrease of the sum normalized L_1 distance function for 60 runs of 1000 iterations

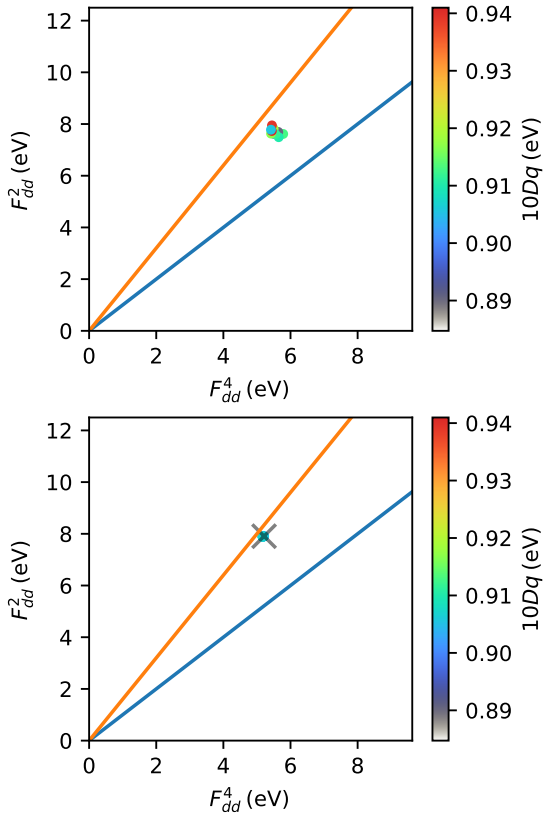

FIG. S2. Top: distribution of 11 GPR evaluations with $d \leq 1.2\chi^2_{\mathrm{min,GPR}}$ for $NiCl_2$. Bottom: results of subsequent greedy optimization starting from the 11 best GPR points. The chosen point is denoted by a gray cross (×).

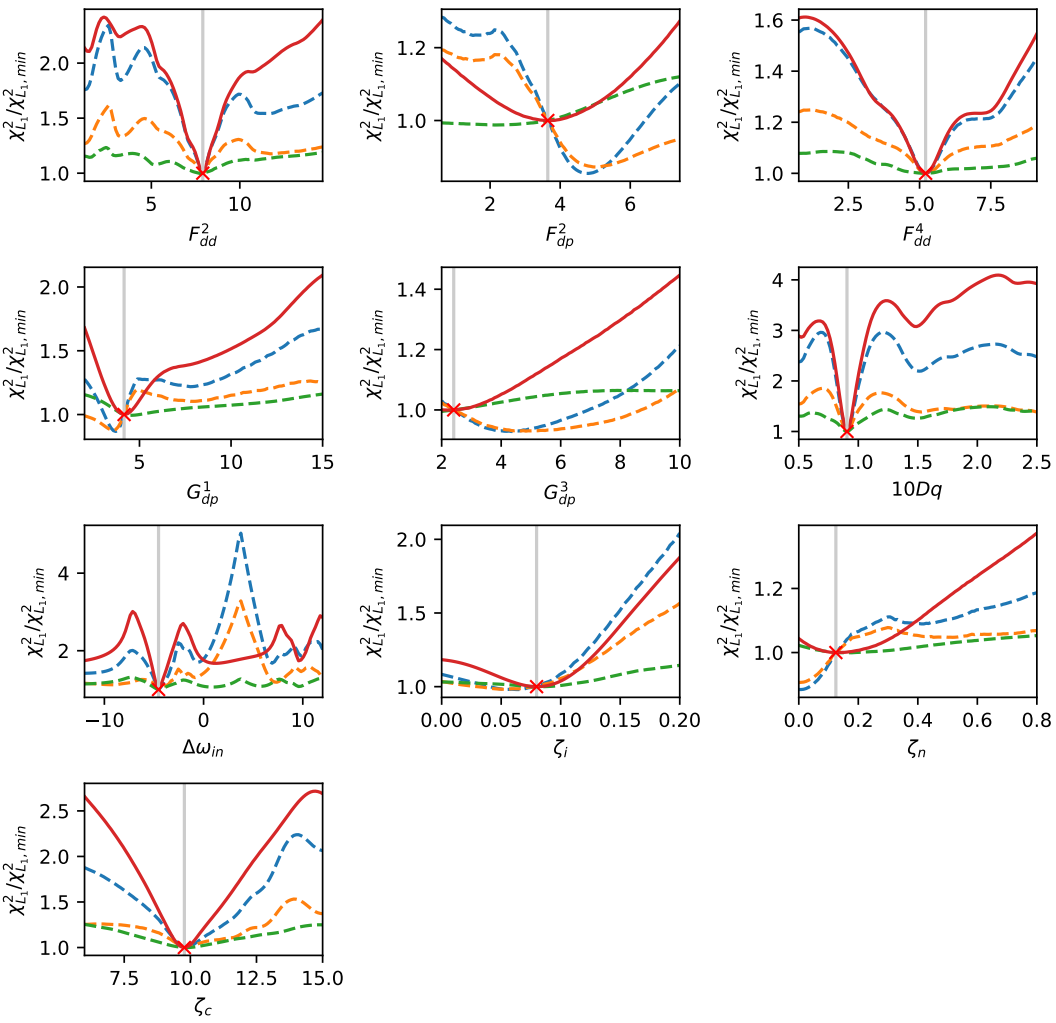

FIG. S3. $NiCl_2$: The behavior of various distance measures around the fine-tuned minimum. Solid red: L1 sum normalized, dashed blue: L1 maximum normalized, dashed orange: L2 sum normalized, dashed green: magnitude of gradient, maximum normalized The fit is highly sensitive to initial Slater and crystal field parameters but weakly depend on intermediate state parameters, especially spin-orbit couplings.

B. Fe_2O_3

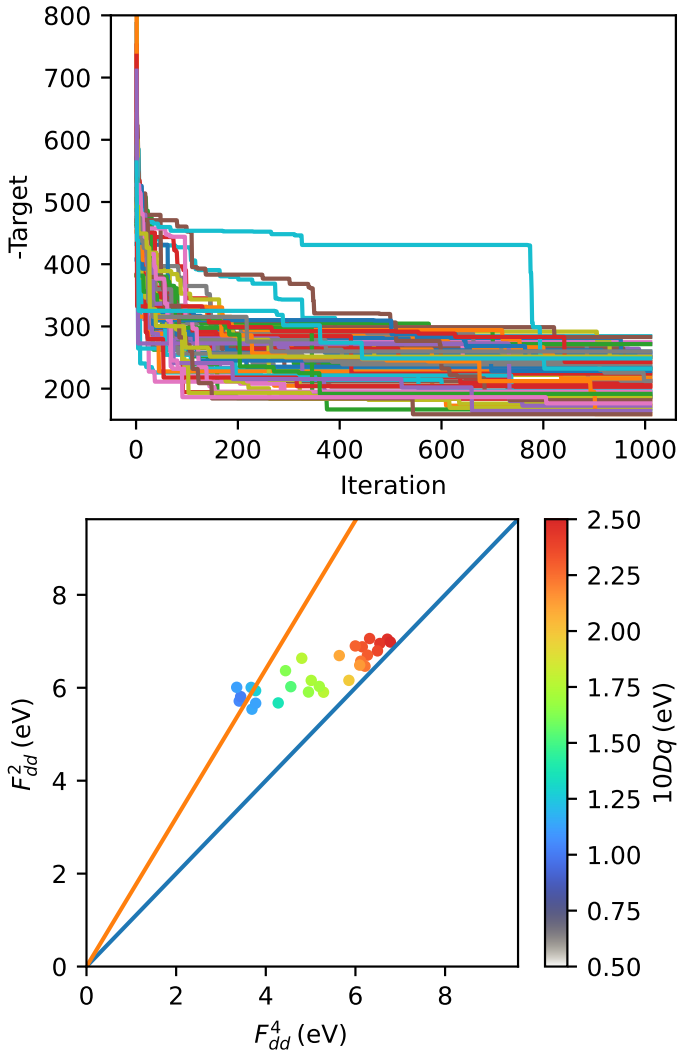

FIG. S4. Top: Decrease of the maximum normalized L1 distance function for Fe_2O_3 over 60 runs of 1000 iterations. Bottom: distribution of the best 28 GPR evaluations serving as starting point of the greedy refinement.

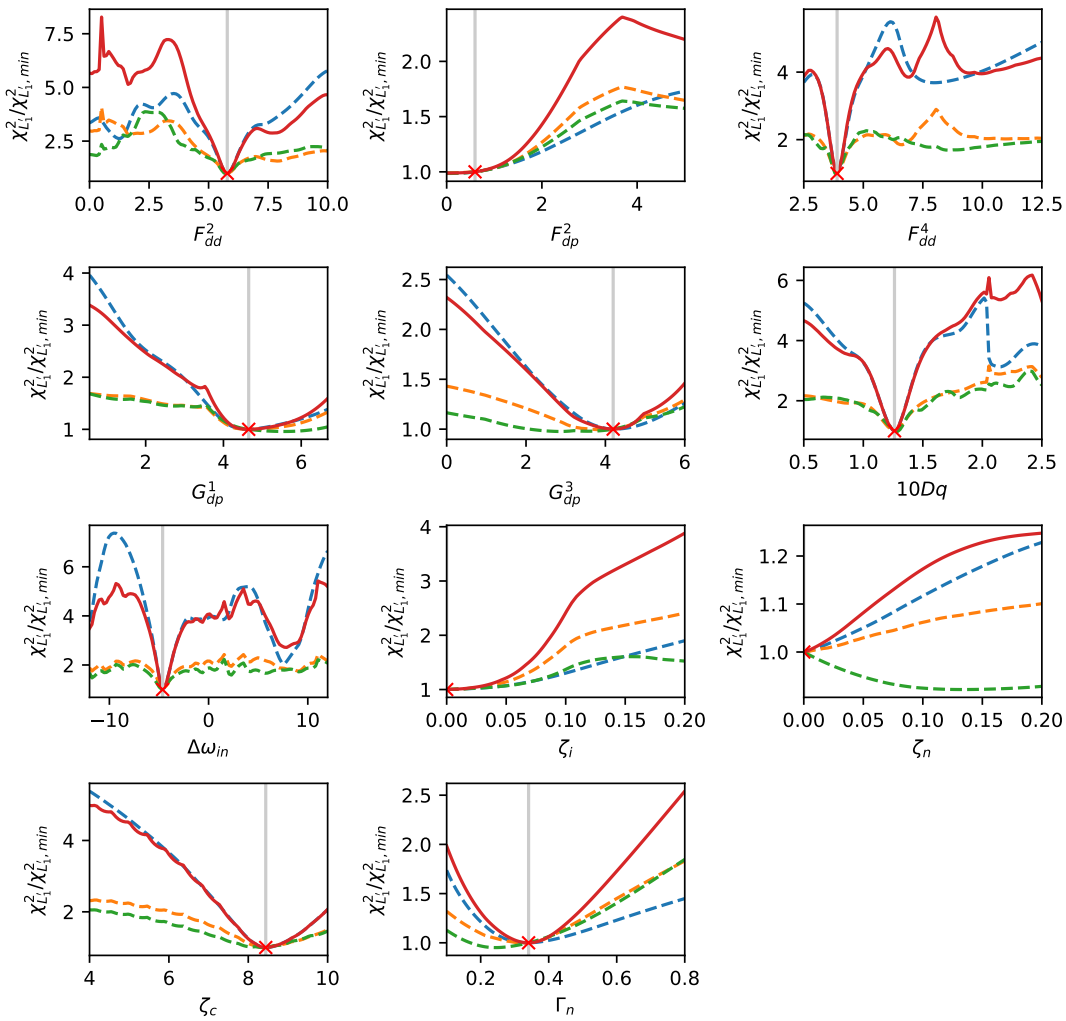

FIG. S5. Fe_2O_3: The behavior of various distance measures around the fine-tuned minimum. Solid red: L1 *maximum* normalized, dashed blue: L1 sum normalized, dashed orange: L2 sum normalized, dashed green: magnitude of gradient, maximum normalized. The fit is highly sensitive to initial Slater and crystal field parameters but has a weaker dependency on intermediate state parameters, especially the spin-orbit coupling ζ_n.

C. Ca_3LiOsO_6

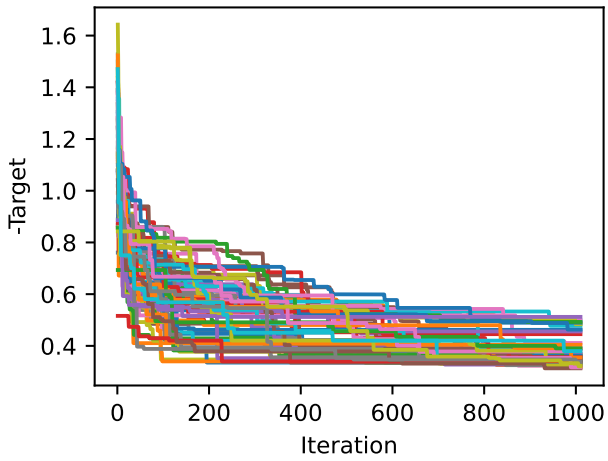

FIG. S6. Ca_3LiOsO_6: Decrease of the sum normalized L1 distance function for 60 runs of 1000 iterations

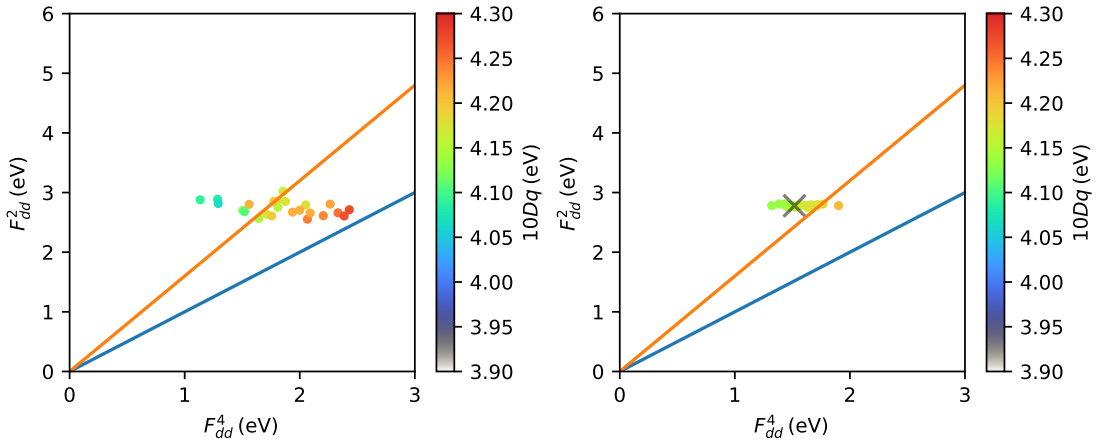

FIG. S7. Left: distribution of 24 GPR evaluations with $d \leq 1.06\chi^2_{\mathrm{min,GPR}}$ for Ca_3LiOsO_6. Right: results of subsequent greedy optimization starting from the 24 best GPR points. The final accepted point is denoted by a gray cross (×).

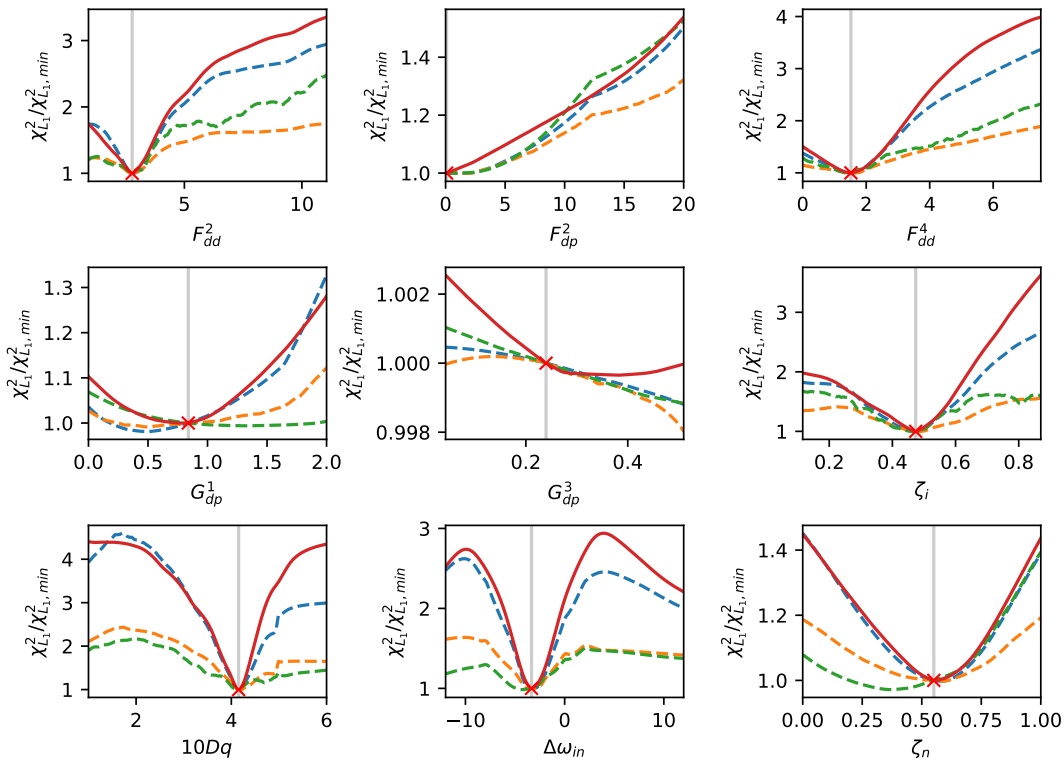

FIG. S8. Ca_3LiOsO_6: The behavior of various distance measures around the fine-tuned minimum. Solid red: L1 sum normalized, dashed blue: L1 maximum normalized, dashed orange: L2 sum normalized, dashed green: magnitude of gradient, maximum normalized. The fit is highly sensitive to initial Slater and crystal field parameters but has a weaker dependency on intermediate state parameters.

[1] M. Dresselhaus, G. Dresselhaus, and A. Jorio, *Group Theory. Application to the Physics of Condensed Matter* (Springer, 2008).