

Hamiltonian parameter inference from RIXS spectra 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: August 4, 2025)

Identifying model Hamiltonians is a vital step toward creating predictive models of materials. Here, we combine Bayesian optimization with the EDRIXS numerical package to infer Hamiltonian parameters from resonant inelastic X-ray scattering (RIXS) spectra within the single atom approximation. To evaluate the efficacy of our method, we test it on experimental RIXS spectra of NiPS₃, NiCl₂, Ca₃LiOsO₆, and Fe₂O₃, and demonstrate that it can reproduce results obtained from hand-fitted parameters to a precision similar to expert human analysis while providing a more systematic mapping of parameter space. Our work provides a key first step toward solving the inverse scattering problem to extract effective multi-orbital models from information-dense RIXS measurements, which can be applied to a host of quantum materials. We also propose atomic model parameter sets for two materials, Ca₃LiOsO₆ and Fe₂O₃, that were previously missing from the literature.

I. INTRODUCTION

Quantum materials are at the forefront of condensed matter research, due to their rich physics and potential as key components of future technologies [1–4]. These systems are often governed by quantum fluctuations and strong many-body interactions, which are not well described by conventional single-particle theories or *ab initio* methods. As such, a central paradigm for progress in quantum materials research has been to identify minimal effective models [5] that capture a material’s low-energy properties while remaining tractable. Leveraging such models is also a powerful way to understand and predict new material properties and identify broader organizing principles.

Reliably identifying a valid minimal effective model for a given system can be extremely challenging. A paradigmatic example of this is the open question of whether the single-band Hubbard model is the correct low-energy model for high- T_c cuprates [6–11]. Other examples include magnetic van der Waals (vdW) materials like NiPS₃ and NiI₂, which host novel many-body excitons that depend crucially on the magnetic state of the lattice [12–14]. Despite extensive studies, the electronic character, mobility, and magnetic interactions of these excitons are still being debated [12, 13, 15], in part because their low-energy Hamiltonian has yet to be conclusively identified.

Theoretically, one can attempt to derive an effective low-energy model from a high-energy Hamiltonian

by integrating out various degrees of freedom. However, strongly correlated systems often exhibit a near-degeneracy of low-energy states that are easily affected by perturbing interactions. This situation can easily bias results if small-but-relevant interactions are not included in the downfolding process. Alternatively, one can attempt to derive a low-energy model using perturbation theory; however, in some cases, the resulting effective couplings can become non-analytic functions of the high-energy model parameters [16].

An alternative line of attack is to take a data-driven approach in the so-called “inverse scattering problem,” where one attempts to extract the correct effective low-energy Hamiltonian from spectroscopic measurements. This approach has been bolstered by recent advances in instrumentation for spectroscopic techniques such as photoemission and inelastic neutron, electron, and photon scattering spectroscopies, as well as the development of new many-body algorithms. The former allows researchers to collect a large amount of high-quality data for a material in a short time, while the latter provides new capabilities for predicting the spectra for a proposed Hamiltonian (the “direct scattering problem”). In real experiments, the inverse scattering problem is particularly relevant as model parameters are not known a priori, and obtaining them from first-principles calculations is often a challenging, if not impossible, task. Today, many practitioners of inverse scattering will hand-tune the parameters of the model to match experimental data. Fitting data by hand is laborious, requires considerable expertise, and can be error prone since humans can only visually inspect a rather modest number of candidate solutions.

The past few decades witnessed rapid growth of available computational resources, accompanied by the swift development of sophisticated algorithms to solve the di-

* mlajer@bnl.gov

† sjohn145@utk.edu

‡ mdean@bnl.gov

rect scattering problem. At the same time, advances in machine learning have had a tremendous impact on the materials science community [17–20]. In particular, new insights from machine learning approaches [16, 17, 20–32] have made it possible to attack the inverse scattering problem from a purely data-driven framework. Inelastic neutron scattering (INS) provides a suitable testing ground for developing this approach, as its cross-section is well understood, and highly efficient direct solvers are available for a broad class of materials [33–36].

Since INS is very sensitive to magnetic excitations, the above-mentioned experimental efforts have so far primarily focused on extracting low-energy spin Hamiltonians for magnetic insulators [37–39]. Our work extends these efforts to resonant inelastic X-ray scattering (RIXS) data, which allows us to access a much broader range of materials and Hamiltonians. RIXS is a photon-in photon-out spectroscopic method in which the energy, momentum, and polarization of an incoming x-ray photon are transferred to a material’s intrinsic momentum-resolved spin, charge, orbital, and lattice excitations [4]. Due to its resonant nature, RIXS can study small samples, including monolayers, while still maintaining bulk sensitivity. RIXS is also suitable for monitoring ultra-fast responses of material in pump-probe experiments [40, 41].

RIXS experiments have access to a large energy window spanning from ≈ 0.02 –10 eV and thus provide access to both high- and low-energy sectors of a given material. This aspect makes it an ideal tool for the inverse scattering problem as it can be used to extract effective models for different energy scales. Moreover, methods for computing RIXS spectra for correlated systems have achieved a degree of standardization and become widely available with packages like EDRIXS [42] and Quanty [43]. Given the maturity of the experimental and numerical methods, the time is right to begin developing systematic approaches for the RIXS inverse scattering problem. Here, we undertake such an effort by employing Bayesian optimization with EDRIXS, an exact diagonalization-based open-source RIXS solver [42], to predict model Hamiltonian parameters from experimentally measured spectra. We focus on the single ion model, which can provide quantitative descriptions of dd excitations in transition metal complexes. We demonstrate its power on experimental spectra of NiPS_3 [13], NiCl_2 [14], Fe_2O_3 [44], and $\text{Ca}_3\text{LiOsO}_6$ [45]. The resulting parameters are then used to annotate and predict the properties of experimentally observed peaks, including their dependence on temperature and polarization. This work provides the foundation for a fully automated solution to the RIXS inverse scattering problem.

The paper is organized as follows: Sec. II A describes the atomic approximation used to model the RIXS spectra. Sec. II B introduces the active learning techniques used to solve the inverse scattering problem. Sec. III describes how our methods combine the techniques above to infer Hamiltonians from RIXS data. Sec. IV presents results for several compounds that have been studied pre-

viously in the literature. Finally, Sec. V provides some concluding remarks and discussion.

II. METHODS

A. RIXS Modeling

We compute RIXS using the EDRIXS package [46, 47]. This software implements the Kramers-Heisenberg (KH) equation, which is the result of treating the photon-matter interaction using second-order perturbation theory.¹ Denoting the momentum, energy, and polarization of the incoming (outgoing) X-rays as $\hbar\mathbf{k}$, $\hbar\omega_{\mathbf{k}}$, and $\hat{\epsilon}$ ($\hbar\mathbf{k}'$, $\hbar\omega_{\mathbf{k}'}$, and $\hat{\epsilon}'$), respectively, the intensity for RIXS can be written as

$$I_{\epsilon\epsilon'}(\hbar\omega_{\mathbf{k}}, \hbar\omega_{\mathbf{k}'}, \mathbf{k}, \mathbf{k}'; T) \propto \frac{1}{\mathcal{Z}(T)} \sum_i e^{-E_i/(k_B T)} \times \sum_f |M_{fi}|^2 \delta(E_f + \hbar\omega_{\mathbf{k}'} - E_i - \hbar\omega_{\mathbf{k}}). \quad (1)$$

Here, M_{fi} represents the matrix element from the system’s initial state i with energy E_i to its final state f with energy E_f via an intermediate state with a core hole, and $\mathcal{Z}(T) = \sum_i e^{-E_i/(k_B T)}$ is the partition function at temperature T . k_B is the Boltzmann constant.

The examples considered in this work correspond to L -edge RIXS, which involves a d -electron valence state and a $2p$ core hole. We wish to compute the so-called dd excitations, where the final states entail a reconfiguration of electrons within the $3d$ manifold, and which manifest as peaks at photon energy loss $E_{\text{loss}} = \hbar(\omega_{\mathbf{k}} - \omega_{\mathbf{k}'})$. The matrix element for RIXS within the KH formalism is given by

$$M_{fi} = \sum_n \frac{\langle f | \mathcal{D}_{\mathbf{k}'\hat{\epsilon}'}^\dagger | n \rangle \langle n | \mathcal{D}_{\mathbf{k}\hat{\epsilon}} | i \rangle}{E_n - E_i - \hbar\omega_{\mathbf{k}} + i\Gamma_n/2}, \quad (2)$$

where $\Gamma_n/2$ is the inverse core-hole lifetime in units of energy and E_n is the energy of the intermediate state. $\mathcal{D}_{\mathbf{k}\hat{\epsilon}}$ is the operator describing the absorption of a phonon, promoting a core electron from a $2p$ state into the d valence states. Similarly, $\mathcal{D}_{\mathbf{k}'\hat{\epsilon}'}^\dagger$ describes the photon emission process via a d to $2p$ transition. $|i\rangle$ and $|f\rangle$ are the eigenstates of the initial state Hamiltonian \hat{H}_i , whereas $|n\rangle$ represents the eigenstates of the intermediate state Hamiltonian \hat{H}_n with a core hole. Due to the attractive potential caused by the core hole, the cross-section is dominated by local transitions, so we take the common approximation of treating the process in the atomic limit, where transitions occur within effective local $3d$ orbitals [4].

¹ For a detailed derivation of this expression from the light-matter interaction, we refer the reader to the review by Ament *et al.* [48].

Adopting the second quantization formalism, \hat{H}_i and \hat{H}_n take the general form

$$\hat{H} = \sum_{\alpha\beta} t_{\alpha\beta} \hat{f}_{\alpha}^{\dagger} \hat{f}_{\beta} + \sum_{\alpha\beta\gamma\delta} U_{\alpha\beta\gamma\delta} \hat{f}_{\alpha}^{\dagger} \hat{f}_{\beta}^{\dagger} \hat{f}_{\delta} \hat{f}_{\gamma}, \quad (3)$$

where the indices run through atomic valence and core orbitals and $\hat{f}_{\alpha}^{\dagger}$ creates an electron in spin-orbital α . The first term of Eq. (3) includes the crystal field as well as spin-orbit coupling and, if appropriate, an applied magnetic field. Here we consider the cubic case in which $t_{\alpha\beta}$ is diagonal in the basis of *real* spherical harmonics: it has eigenvalues $6D_q$ in the e_g orbitals ($d_{x^2-y^2}$ and d_{z^2}) and $-4D_q$ for the t_{2g} orbitals (d_{xy} , d_{xz} and d_{yz}). Valence band spin-orbit coupling for the d states in the initial (intermediate) state is parameterized by ζ_i (ζ_n). The core hole spin-orbit coupling is parameterized by ζ_c .

The second term of Eq. (3) accounts for the Coulomb interactions with the d shell and between the p and d shells and is parameterized by Slater integrals [49]. The relevant terms for the initial Hamiltonian are F_{dd}^2 and F_{dd}^4 . In the intermediate state, these are supplemented by additional parameters describing the interactions between the valence electrons and the core-hole, F_{dp}^2 , F_{dp}^4 , G_{dp}^1 , and G_{dp}^3 . For an atomic model, as applied here, the parameters F_{dd}^0 and F_{dp}^0 only produce overall shifts in the eigenenergies. In general, the valence Slater integrals can also be modified in the presence of the core hole, but we make the approximation of F_{dd}^2 and F_{dd}^4 being the same in the initial and intermediate states.

Since the X-ray wavelength tends to be larger than the extent of the atomic orbitals, we treat $\mathcal{D}_{\mathbf{k}\hat{\epsilon}}$ and $\mathcal{D}_{\mathbf{k}'\hat{\epsilon}'}$ within the dipole approximation. In the atomic limit this means that $\mathcal{D}_{\mathbf{k}\hat{\epsilon}} = \sum_{\alpha,\beta} \langle \phi_{\alpha} | \hat{\epsilon} \cdot \hat{r} | \phi_{\beta} \rangle \hat{f}_{\alpha}^{\dagger} \hat{f}_{\beta}$, where β indexes a core spin-orbital and α is a valence spin-orbital and $|\phi_{\alpha(\beta)}\rangle$ are the corresponding atomic orbital states.

To account for the finite experimental energy resolution and the finite lifetime of the dd excitations, the delta function in Eq. (1) is represented by a Lorentzian with full-width at half-maximum (FWHM) γ . Both γ and Γ_n are considered as constants independent of the state in question. Because the absolute core-hole energy is undefined in this model, we introduce an adjustable energy offset x_{offset} to align the theoretical spectra with the experimental data.

To interpret the RIXS spectra and distinguish different classes of solutions to the inverse scattering problem, it is helpful to be able to label the different n -electron eigenvectors of \hat{H}_i . In the absence of spin-orbit coupling (SOC) and crystal field splitting, the eigenvalues of the spin and orbital angular momentum operators are good quantum numbers and form useful labels. This breaks down for the more general case. Still, the eigenvectors can be effectively annotated by the states that transform according to a certain irreducible representations of the octahedral group, as illustrated in Fig. 1. When the octahedral symmetry is not exact, we use the weights of the eigenvectors within definite symmetry sectors to

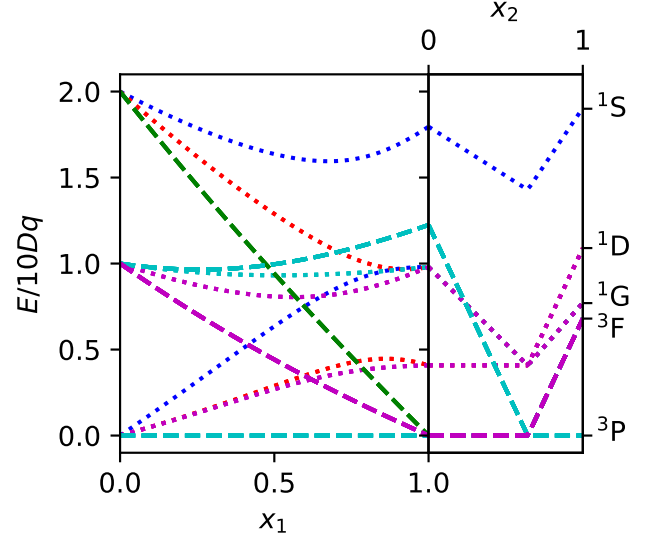


FIG. 1. Tanabe-Sugano-style plots for d -shell electronic (initial) Hamiltonians of the form $x_1 U_{\text{coul}}^{F^2} + (1-x_1) V_{\text{cf}} + x_2 U_{\text{coul}}^{F^4} - (1-x_2) U_{\text{coul}}^{F^2}$ with occupation numbers 2 or 8. On the left subplot, $x_2 = 0$ and x_1 is varied (x_1 subplot). On the right subplot, $x_1 = 1$ and x_2 is varied. Dotted lines correspond to spin singlets, dashed lines denote triplets. States transforming according to different irreducible representations of the octahedral group are shown in different colors: A_1 – blue, A_2 – green, E – red, T_1 – teal, T_2 – purple. Ground state energies are subtracted so that occupations 2 and 8 yield the same figure. This plot gives an overview of the spectrum of H_i as a function of its three most important parameters F^2 , F^4 , and $10Dq$, with other parameters fixed to zero.

approximately characterize the states. Our annotation method is described in Sec. S1 of the Supplemental Material (SM) [50].

B. Bayesian optimization via Gaussian process regression

The inverse scattering problem involves searching a high-dimensional parameter space to identify the model Hamiltonian that best reproduces experimental spectra. Since solving the direct problem (i.e., computing the theoretical RIXS spectrum from a given Hamiltonian) is potentially computationally expensive, it is imperative to minimize the number of function evaluations required during inference. To address this, we employ a Bayesian optimization strategy that builds a surrogate model of the distance function between simulated and experimental spectra, guiding the search for optimal parameters in a sample-efficient manner [51].

1. Distance metrics

To compare theoretical and experimental spectra, we define a distance function χ^2 operating directly on the two-dimensional (2D) spectral images. In the following formulae, we denote the entries of the 2D arrays containing the experimental (reference) and simulated spectral intensities by R_{ij} and S_{ij} , respectively. The summation over indices i, j, k, l , etc. spans over each row and column of the corresponding arrays. Several pixel-wise metrics are considered. They include the sum normalized L_p distance

$$\chi_{L_p}^2 = \left(\sum_{ij} |r_{ij} - s_{ij}|^p \right)^{\frac{1}{p}}, \quad (4)$$

where

$$r_{ij} = \frac{R_{ij}}{\sum_{kl} R_{kl}} \quad s_{ij} = \frac{S_{ij}}{\sum_{kl} S_{kl}};$$

the maximum normalized L_p distance

$$\chi_{L'_p}^2 = \left(\sum_{ij} \left| \frac{R_{ij}}{\max R} - \frac{S_{ij}}{\max S} \right|^p \right)^{\frac{1}{p}}, \quad (5)$$

and the numerical gradient

$$\chi_g^2 = \sum_{ij} \left| \sqrt{\left(\frac{r_{ij} - r_{i,j-1}}{\Delta\omega_{\text{in}}} \right)^2 + \left(\frac{r_{ij} - r_{i-1,j}}{\Delta E_{\text{loss}}} \right)^2} - (r \leftrightarrow s) \right|, \quad (6)$$

which emphasizes features like sharp peaks or other abrupt changes in the spectrum. Here, $\Delta\omega_{\text{in}}$ and ΔE_{loss} are the energy spacings along the ω_{in} and E_{loss} axes, respectively.

2. Gaussian process regression

Our approach is based on iteratively constructing and refining a probabilistic model of the distance landscape, defined over the space of Hamiltonian parameters. The central idea is to treat the evaluation of the RIXS spectrum as a black-box function that is expensive to query but smooth and continuous over physically relevant regions.

To model this function, we use Gaussian process regression (GPR), a flexible, non-parametric regression method that defines a distribution over functions [52, 53]. In our case, the input to the GPR is a vector of Hamiltonian parameters \mathbf{z} , such as crystal field splitting $10D_q$, Slater integrals, etc. The output is the value of a distance function χ^2 , which quantifies the discrepancy between the RIXS spectrum generated by those parameters

and the experimental spectrum. Formally, GPR models the function $\chi^2 : \mathbb{R}^n \rightarrow \mathbb{R}$, where n is the number of free parameters being optimized.

A Gaussian process is fully specified by a mean function $f_{\text{prior}}(\mathbf{z})$ and a kernel function $k(\mathbf{z}_1, \mathbf{z}_2)$. Given a set of previously evaluated parameter points and their corresponding distance values, a Gaussian process provides a posterior distribution over functions that can be used to predict the distance $f_{\text{reconst}}(\mathbf{z})$ at unseen parameter configurations

$$f_{\text{reconst}}(\mathbf{z}) = f_{\text{prior}}(\mathbf{z}) + \sum_{ij} k(\mathbf{z}, \mathbf{z}_i) (\mathbf{K}^{-1})_{ij} [f(\mathbf{z}_j) - f_{\text{prior}}(\mathbf{z}_j)], \quad (7)$$

along with a confidence interval

$$\delta f_{\text{reconst}}(\mathbf{z}) = k(\mathbf{z}, \mathbf{z}) - \sum_{ij} k(\mathbf{z}, \mathbf{z}_i) (\mathbf{K}^{-1})_{ij} k(\mathbf{z}_j, \mathbf{z}), \quad (8)$$

where the covariance matrix is

$$\mathbf{K} = \begin{pmatrix} k(\mathbf{z}_1, \mathbf{z}_1) & k(\mathbf{z}_1, \mathbf{z}_2) & \dots & k(\mathbf{z}_1, \mathbf{z}_n) \\ k(\mathbf{z}_2, \mathbf{z}_1) & \ddots & & \\ \vdots & & \ddots & \\ k(\mathbf{z}_n, \mathbf{z}_1) & \dots & & k(\mathbf{z}_n, \mathbf{z}_n) \end{pmatrix}. \quad (9)$$

Throughout this work, we use the Matérn kernel with $\nu = \frac{5}{2}$

$$k(\mathbf{z}_1, \mathbf{z}_2) = \left(1 + \frac{\sqrt{5}}{l} d(\mathbf{z}_1, \mathbf{z}_2) + \frac{5}{3l} d(\mathbf{z}_1, \mathbf{z}_2)^2 \right) \times \exp \left(-\frac{\sqrt{5}}{l} d(\mathbf{z}_1, \mathbf{z}_2) \right), \quad (10)$$

where the length-scale parameter l is optimized internally through the regression process. The Bayesian prior $f_{\text{prior}}(\mathbf{z})$ is set to zero in the following.

3. Acquisition function

The acquisition function is a central component in active learning, which determines the next sampling point in the parameter space based on the surrogate model constructed by GPR. This function quantifies the trade-off between exploration of regions with large uncertainties and exploitation of regions predicted to yield low distances, i.e., good descriptions of the experimental spectra. In our approach, we employ the Upper Confidence Bound (UCB) acquisition function, which balances the mean prediction of the GPR and its associated uncertainty [54]. It is defined as:

$$\text{UCB}(\mathbf{z}) = f_{\text{reconst}}(\mathbf{z}) + \kappa \delta f_{\text{reconst}}(\mathbf{z}), \quad (11)$$

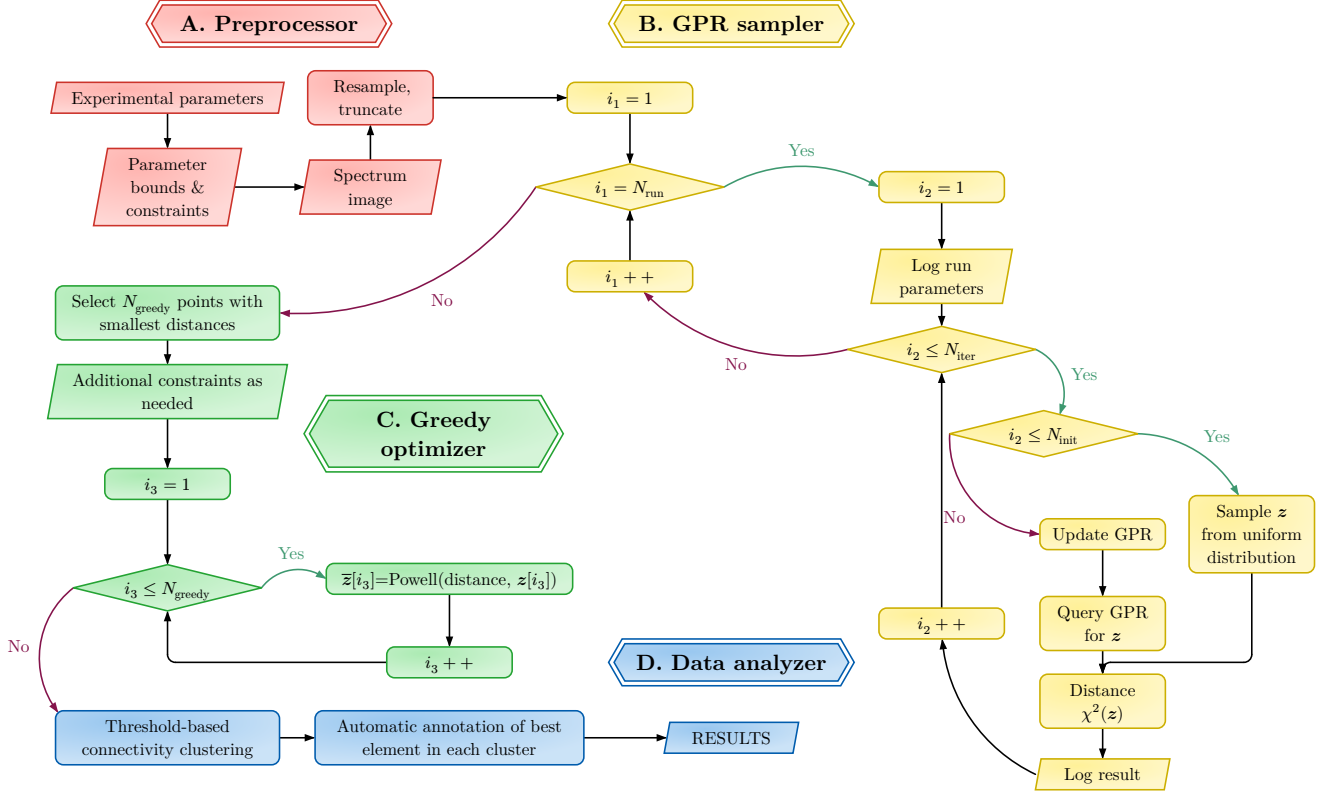


FIG. 2. Flowchart of the Bayesian optimization algorithm used to solve the inverse scattering problem for RIXS. A. The preprocessor imports and processes a spectrum image from a RIXS experiment. B. The GPR sampler builds a set of N_{run} models of the distance function via an active learning protocol. Each run consists of N_{iter} GPR iterations, followed by N_{init} initial evaluations at random points. C. Parameters obtained from the N_{greedy} smallest queried distances are refined by a subsequent greedy optimizer. D. Finally, the results are presented by the data analyzer.

where $f_{\text{reconst}}(\mathbf{z})$ is the Gaussian process's predicted distance at parameter configuration \mathbf{z} , $\delta f_{\text{reconst}}(\mathbf{z})$ is the model's uncertainty estimate, and κ is a hyperparameter controlling the exploration-exploitation balance. A larger κ encourages exploration by favoring points with high uncertainty, while a lower value promotes exploitation by prioritizing low predicted distances. The UCB acquisition function enables efficient navigation of the high-dimensional Hamiltonian parameter space by guiding the sampling toward informative regions. This aspect is crucial for minimizing the number of RIXS simulations during the optimization process.

III. IMPLEMENTATION OF THE INVERSE SCATTERING PROBLEM

Figure 2 presents a flowchart of the method. We employ a two-stage optimization strategy that combines global exploration with local refinement to efficiently minimize the distance between simulated and experimental RIXS spectra in a high-dimensional parameter space.

In the first stage, the experimental data is prepro-

cessed. In the second stage, we use GPR to construct a surrogate model of the distance function over the Hamiltonian parameter space. The GPR is initialized with a small number of randomly sampled points, and subsequent evaluations are guided by the acquisition function Eq. (11). To improve efficiency and stability, we restrict optimization at this stage to a subset of parameters that have the strongest influence on the RIXS spectra, while holding others such as energy loss broadening γ and most spin-orbit couplings fixed. This design choice helps reduce dimensionality and avoids uninformative regions of parameter space. Further details can be found in Sec. III B.

In the third stage, we apply a local, derivative-free optimizer (in our case, Powell's method via `py-BOBYQA`) to refine the best candidates found by GPR [55]. This step not only improves the precision of previously optimized parameters but also reintroduces variations for any parameters fixed in the second stage, enabling a full-space optimization. This refinement enhances convergence and consistently leads to lower distance metrics.

Finally, in the fourth stage, the results are processed, peak features are classified according to their symmetry,

and the results are output.

In the following sections, we provide a detailed description of each stage of the process.

A. Preprocessing

We take the experimental spectrum in the form of a two-dimensional array of intensities interpolated onto a regular grid. The array indices correspond to different values of the incident energy ω_{in} and energy loss E_{loss} . The intensities are positive, but given only up to an overall multiplicative constant. The experimental spectrum typically exhibits many peaks, some of which are better described by the single ion model than others. Since we focus on dd -excitations, the first step is to crop the spectrum in the energy loss direction, omitting the effect of the elastic peak, low-energy collective excitations, and excitations to other high-energy orbitals (e.g., charge transfer excitations). Restricting the spectrum to the relevant energy range of the model used is crucial and significantly impacts the reliability of the results.

B. GPR sampler

Once preprocessing is complete, we provide the processed spectrum to the GPR sampler, which constructs a model of a distance function χ^2 as a function of the model parameters, \mathbf{z} . Evaluating the distance function involves running the direct scattering problem, that is using EDRIXS to compute the RIXS spectra.

We have found that it is beneficial to use the sum-normalized L_1 distance $\chi_{L_1}^2$ when the RIXS spectrum is not too crowded, e.g., away from a half-filled initial valence shell. This distance function normalizes with respect to the volume under the spectrum, so it gives considerable weight to fainter side peaks that the max-norm measure $\chi_{L_1}^2$ is more prone to missing. In contrast, when the initial configuration is closer to half-filling, the parameter space consists of a plethora of configurations with intensities widely spread in energy space. This structure leads to a rich and shallow landscape for the sum norm, which makes optimization difficult. Under these circumstances, we have found that the max norm, which gives extra emphasis to the region around the brightest peak, is more robust. We have also found that the gradient norm generally underperformed the sum and max norms. The behaviors of each norm around the final parameter set for each material are reported in Fig. 9 of App. A and Sec. S2 of the SM [50].

To build the Gaussian approximation of χ^2 , we initially evaluate the distance function at a small number $N_{\text{init}} = 10$ of random points in parameter space. The Gaussian process provides an estimate χ^2 as well as its uncertainty between the calculated points. We then feed this information to the acquisition function [see Eq. (11)] and maximize this function to yield the next queried

point (at which χ^2 is to be evaluated). Once the value of χ^2 is known at the new point, the GPR model is updated, and the acquisition is maximized again to obtain the next point. The process is iterated for $N_{\text{iter}} = 1000$ iterations. This active learning approach is intended to aid the solution of the global minimization of the distance function with a minimal number of expensive spectral simulations. To improve the results and get a sense of the robustness of the optimization, we repeat the buildup of the Gaussian model from different initial points for N_{run} times in total. This yields a total number $N_{\text{run}} \times (N_{\text{init}} + N_{\text{iter}})$ evaluations of χ^2 (i.e., calls to the EDRIXS solver).

Since the number of a priori floating parameters is rather large, we chose a subset of parameters with respect to which χ^2 is optimized by GPR. This selection is mainly dictated by the sensitivity of the distance function against varying the corresponding parameter. Such dependencies can be predicted theoretically and confirmed numerically, as detailed in Fig. 9 and Figs. S3, S5, and S8 of the SM [50]. Another concern pertaining to the broadening parameters is that floating them often allows the optimization to enter uninteresting valleys of χ^2 where the broadening gets very large and the distinguishing features of the RIXS spectra are washed out. To avoid this, we fix the value of the core-hole broadening $\Gamma/2$ at this stage, as well as (with the exception of $\text{Ca}_3\text{LiOsO}_6$) all spin-orbit couplings to their atomic values. We also set the final state broadening γ empirically to match the typical width of the dd excitations. GPR then focuses on fitting $10Dq$, the Slater parameters, the offset x_{offset} , and in the case of $\text{Ca}_3\text{LiOsO}_6$, the initial and intermediate state valence spin-orbit couplings with the temporary constraint $\zeta_{v,i} = \zeta_{v,n}$.

The results of different runs are combined and ranked by distance function value before being passed to a subsequent greedy optimizer. In doing this, one tests for the presence of multiple local minima in the distance function and the extent to which these have similar fit quality. This stage further restricts the parameter space to the most interesting candidate regions and excludes any obviously unphysical local minima identified by GPR. For example, one can exclude small-distance regions where $F_{dd}^2 \ll F_{dd}^4$.

Although we found that GPR itself is efficient in locating promising regions in parameter space, we found it advantageous to add a further optimization step to refine further the parameters obtained.

C. Greedy optimizer

Since the GPR sampling balances exploration and optimization in a complex parameter space, it is not necessarily best suited for fine optimization when close to a minimum in the distance landscape. In principle, one could fine tune the fitting process by adjusting the hyperparameter κ controlling the balance between exploration and exploration. However, we have found it more efficient

to instead include a step in which we collect the “best” $N_{\text{greedy}} \sim 10 - 30$ evaluated points with the smallest distance function values from the candidate parameter regions. These points are used as initial points to start a greedy optimization of the distance function. The greedy optimizer uses a refined version of Powell’s method, provided by the package `py-BOBYQA` [55]. At this stage, the optimizer extends the set of fitted parameters to now include the spin-orbit couplings and other parameters that were omitted from GPR. The optimization started from a particular initial point typically converges in a few thousand function evaluations.

Besides improving on the parameter estimates, the greedy optimization provides valuable information on the performance of the GPR minimization and the robustness of the method in general. The greedy optimizer is used to obtain uncertainty intervals for the local solutions.

D. Data analyzer

The greedy-optimized set of points is grouped based on a threshold-based connectivity clustering approach. (To illustrate the kinds of differences that are typically found in these clustered solutions, Fig. 11 of App. A plots the spectra obtained from the best fits of three closely packed clusters obtained for Fe_2O_3 .) Our analysis pipeline presents the user with a list of physically distinct solutions ordered by quality of fit. The robustness of the method is benchmarked by asking the question, how many of the runs provide points that are relaxed into the peak corresponding to the accepted fit. We use the N_{greedy} best points of the GPR sampler as a proxy to estimate the robustness of the fitting procedure.

Finally, the eigenstates of the initial Hamiltonian for the best fit are annotated by eigenvalues of a set of approximately conserved quantities and discrete symmetry labels. We developed our own code to perform this annotation automatically, which is detailed in Section S1 in the SM [50].

IV. RESULTS FOR EXPERIMENTAL SPECTRA

A. NiPS_3

NiPS_3 is a layered van der Waals crystal in which the active Ni d^8 ions are coordinated by six S atoms in approximately cubic symmetry [56, 57]. It has attracted significant attention for its optoelectronic properties, including a novel magnetic exciton [12, 13] and its two-dimensional magnetism [58]. We consider this material because it represents a challenging case in which the atomic model can capture some, but not all, of the physics of this material in light of its small charge-transfer energy [12, 13].

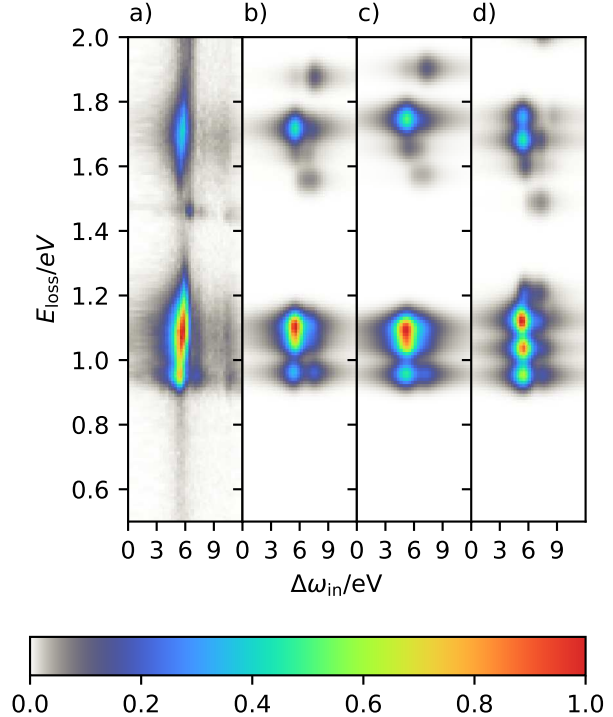


FIG. 3. Experimental and simulated spectra for NiPS_3 . (a) experimental data from Ref. [13], (b) hand-fitted (reference), (c) result of GPR, (d) result after greedy fine-tuning (Powell’s method). The corresponding L_1 sum distances for the models shown in panels b-d are $\chi^2_{L_1} = 0.692, 0.694$, and 0.480 , respectively.

We use the experimental RIXS spectrum from Ref. [13]. The spectrum was measured at the SIX 2-ID beamline of the National Synchrotron Light Source II (NSLS-II) with an energy resolution of 31 meV FWHM. The experimental spectrum was recorded at temperature $T = 40$ K, well below the Néel ordering temperature of 140 K. The geometry involved an incident x-ray angle of $\theta_{\text{in}} = 23^\circ$ and a scattering angle of $2\Theta = 150^\circ$. Intensities were recorded in an incident energy range $\omega_{\text{in}} \in [848.5 \text{ eV}, 857.6 \text{ eV}]$ around the Ni L_3 -edge with π -polarized incident x-rays. The originally measured spectrum was subsequently truncated to the energy loss range $E_{\text{loss}} \in [0.5 \text{ eV}, 2.0 \text{ eV}]$. It was also resampled into an equidistant grid with 40 points in the ω_{in} direction and 151 points in the E_{loss} direction.

1. GPR results

The comparison of experimental spectrum, together with the hand-fit results of Ref. [13] (which we refer to as the “reference” model hereafter), GPR, and greedy-updated simulated spectra are shown in Fig. 3. Fitted parameters after the greedy refinement are reported along

TABLE I. NiPS₃ parameters (in units of eV). The “interval 1.25” and “interval 1.5” refer to the range of values for each parameter that keep the distance function within 1.25× or 1.5× its minimum value, respectively, with all other parameters held fixed.

Parameter	Ref.	Pt. est.	interval 1.25	interval 1.5
F_{dd}^2	5.26	5.50	[5.09, 6.02]	[4.88, 6.53]
F_{dp}^2	6.18	6.70	[2.15, 10.93]	[n/a, 13.17]
F_{dd}^4	3.29	4.50	[3.66, 5.56]	[3.08, 6.49]
G_{dp}^1	2.89	2.74	[-0.43, 4.56]	[-1.81, 5.86]
G_{dp}^3	1.65	2.38	N/A	N/A
$10Dq$	1.07	1.061	[1.02, 1.09]	[1.00, 1.18]
$\Delta\omega_{in}$	N/A	-1.02	[-1.36, -0.65]	[-1.53, -0.42]
ζ_i	0.083	0.123	[0.101, 0.147]	[0.089, 0.162]
ζ_n	0.102	0.0	N/A	N/A
ζ_c	11.2	11.46	[10.78, 12.21]	[10.43, 12.66]
Γ_n	0.6	0.544	[0.273, n/a]	[0.189, n/a]
γ	0.05	0.06	(fixed)	(fixed)

with their reference counterparts in Tab. I. Reference values included in Tab. I are taken from Ref. [13]. The table includes two different uncertainty intervals for each fitted parameter. Intervals 1.25 and 1.5 correspond to the intervals within which the distance function increases by 1.25 and 1.5, respectively, from their minimum values (see also Fig. 9). The distance function $\chi_{L_1}^2$ was used in the optimization.

The results obtained from the GPR and after greedy fine-tuning are in excellent agreement with the reference spectrum, and of compatible (if not superior) to the hand-fitted spectrum. The classification of the excitations is also similar for both the GPR and greedy fine-tuned solutions. The ground state is a triplet 3A_1 state with completely filled t_{2g} subshell and two electrons on the e_g manifold. Here the upper index denotes spin multiplicity and A_1 is the symmetry label with respect to the octahedral group. The lowest lying excitations at around 1–1.2 eV correspond to dd excitations where one electron transfers between the t_{2g} and e_g orbitals. These excitations typically have 3T_2 symmetry, with some mixing of 1E and 3T_1 . The higher energy loss peaks are predominantly 3T_1 with either one or two electrons transferred to e_g orbitals. These annotations align well with previous findings in the literature [13].

It should be noted that the feature around 1.45 eV, which has previously been identified as a Hund’s exciton [13], is not perfectly predicted by the model. This difference, and other small discrepancies between the theory and experiment, arise from the atomic approximation used here. In particular, it has been shown that an Anderson impurity model is required to more accurately capture these excitations due to the rather small energetic difference between the Ni d and S p states in this material [59].

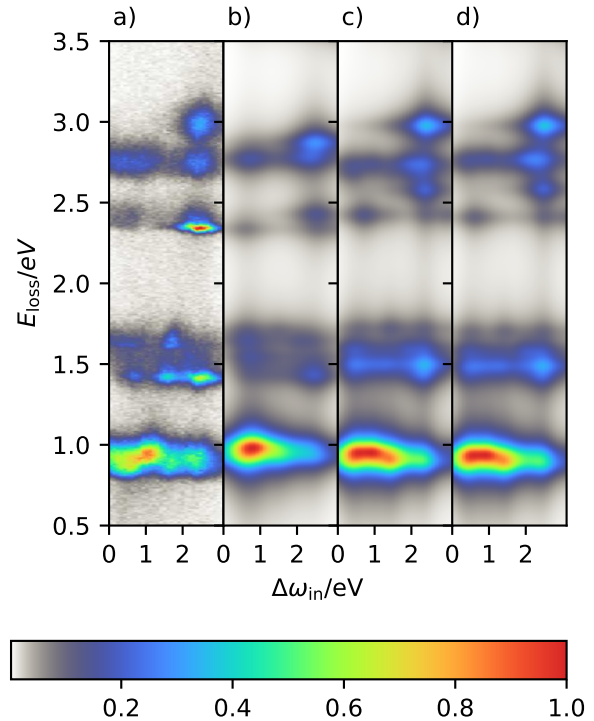


FIG. 4. Experimental and simulated spectra for NiCl₂. (a) experimental data from Ref. [14], (b) hand-fitted (reference), (c) result of GPR, (d) result after greedy fine-tuning (Powell’s method). The corresponding L1 sum distances for panels (b)–(d) are $\chi_{L_1}^2 = 0.496, 0.406$, and 0.373 , respectively.

B. NiCl₂

NiCl₂ is a prototypical van der Waals antiferromagnet and a classic charge-transfer insulator. It adopts a layered rhombohedral structure (space group $R\bar{3}m$) comprised of edge-sharing NiCl₆ octahedra forming a two-dimensional triangular lattice. Each Ni²⁺ ($3d^8$) center is coordinated by six Cl[−] ligands in an approximately octahedral environment.

NiCl₂’s low-lying excitations include extremely sharp spin-singlet dd multiplets ($^1A_{1g}/^1E_g$) stabilized by intra-atomic Hund’s exchange, which are coupled to the ligand environment and exhibit ligand-tuned energies and lifetimes. These aspects make NiCl₂ a benchmark for studying Hund’s excitons, their interaction with lattice and magnetic degrees of freedom, and their propagation (dispersion) in 2D correlated insulators [14].

1. Experimental setup

We use the experimental RIXS spectrum from Ref. [14]. The spectrum was measured at the Soft Inelastic X-Ray (SIX) 2-ID beamline of the NSLS-II with an energy resolution of 31 meV FWHM. The ex-

periment probed the Ni L_3 edge. The spectrum was recorded at temperature $T = 40$ K, with grazing incidence, and the a^* reciprocal-lattice direction lying in the scattering plane. Accordingly, we used scattering angles $\theta_{\text{in}} = 10^\circ$, $2\Theta = 150^\circ$ in the simulation. The incident beam was σ -polarized. Intensities were recorded in the incident energy range of width 3.11 eV, including 32 grid points. The examined energy loss range is $E_{\text{loss}} \in [0.5 \text{ eV}, 3.5 \text{ eV}]$, including 311 grid points in total.

2. GPR results

We show the experimental spectrum along with the reference, GPR, and greedy refined spectra in Fig. 4. Parameter values of the fits are collected in Tab. II. The reference values in Tab. II are extracted from Ref. [14]. Ref. [14] reported a clear double-peak structure at the L_3 edge in the incident energy direction, which is well reproduced by the fits. The ground state electronic configuration was characterized as $^3A_{2g}$ with $t_{2g}^6 e_g^2$ orbital occupation, in complete agreement with our numerics. Ref. [14] observed two sharp excitonic peaks, which they annotated as spin-singlet excitations with $^1A_{1g} / ^1E_g$ symmetry labels. We observe peaks at similar locations, although it is difficult to comment on their relative widths due to the limitations of the single ion model. The candidate peaks are less pronounced than the experimental spectrum. Our peak assignments made for our fitted spectra approximate, but do not completely overlap the peak annotations in Ref. [14]. The peaks at the location of the 1E_g exciton at 1.42 eV energy loss are annotated as 3T_1 . In turn, the 1E_g peak in the simulation mix with two other states with 3T_1 symmetry and are positioned upwards in energy by about 0.1 eV and 0.3 eV, respectively. The other sharp peak of label $^1A_{1g}$ at 2.43 eV is present in the computed spectra as well, albeit shifted upwards by 0.05 eV (GPR) and 0.15 eV (greedy). The other peaks are annotated consistently with Ref. [14], with some variance in the e_g occupation numbers.

C. Fe_2O_3 thin films

We now consider RIXS measurements on $\alpha - \text{Fe}_2\text{O}_3$ (hematite) thin films [44]. This material is a well-known insulating antiferromagnet that hosts Fe^{3+} ($d = 5$) ($S = 5/2$) ions and supports long-distance magnon transport without charge flow – an attractive platform for low-power spintronic devices. In the corundum structure each Fe^{3+} is coordinated by six O^{2-} ligands, forming nearly ideal FeO_6 octahedra. This test case is of particular interest because atomic model fits to Fe L -data data for this material currently do not exist in the literature.

TABLE II. NiCl_2 parameters (in units of eV). The “interval 1.25” and “interval 1.5” refer to the range of values for each parameter that keep the distance function within $1.25\times$ or $1.5\times$ its minimum value, respectively, with all other parameters held fixed.

Parameter	Ref.	Pt. est.	interval 1.25	interval 1.5
F_{dd}^2	7.34 (7.8)	7.91	[7.30, 8.40]	[6.76, 8.91]
F_{dp}^2	4.63	3.65	[n/a, 7.22]	N/A
F_{dd}^4	4.56 (4.85)	5.21	[3.82, 7.56]	[2.33, 8.90]
G_{dp}^1	3.93	4.16	[3.04, 5.69]	[2.43, 9.78]
G_{dp}^3	2.24	2.41	[n/a, 7.31]	N/A
$10Dq$	0.95	0.905	[0.87, 0.94],	[0.86, 0.96]
$\Delta\omega_{\text{in}}$	Pt. est.	-4.54	[-4.88, -4.12],	[-5.11, -3.77]
ζ_i	0.083	0.0797	[n/a, 0.132]	[n/a, 0.16]
ζ_n	0.102	0.125	[n/a, 0.616]	N/A
ζ_c	11.507	9.774	[9.08, 10.61]	[8.64, 11.29],
Γ_n	0.5	0.4125	fixed	fixed

TABLE III. Fe_2O_3 parameters (in units of eV). The “interval 1.25” and “interval 1.5” refer to the range of values for each parameter that keep the distance function within $1.25\times$ or $1.5\times$ its minimum value, respectively, with all other parameters held fixed.

Parameter	Pt. est.	interval 1.25	interval 1.5
F_{dd}^2	5.78	[5.60, 5.99]	[5.50, 6.10]
F_{dp}^2	0.6	[0, 1.67]	[0, 2.11]
F_{dd}^4	3.80	[3.78, 4.03]	[3.71, 4.10]
G_{dp}^1	4.65	[3.97, 6.07]	[3.75, 6.55]
G_{dp}^3	4.20	[2.85, 5.42]	[2.24, n/a]
$10Dq$	1.26	[1.21, 1.31]	[1.18, 1.33]
$\Delta\omega_{\text{in}}$	-4.63	[-4.83, -4.29]	[-4.97, -4.11]
ζ_i	0	[0, 0.0533]	[0, 0.069]
ζ_n	0	N/A	N/A
ζ_c	8.44	[7.99, 9.07]	[7.79, 9.42]
Γ_n	0.34	[0.21, 0.47]	[0.16, 0.54]

1. Experimental setup

We use the experimental Fe L_3 edge spectrum reported in Ref. [44] for this test case. The measurement took place at the SIX 2-ID beamline of the NSLS-II with an energy resolution of 23 meV FWHM. The experimental spectrum was recorded at $T = 100$ K using linearly (π) polarized light and scattering angles $\theta_{\text{in}} = 20^\circ$, $2\Theta = 150^\circ$, and an azimuth angle $\varphi = 0^\circ$. Intensities were recorded in the incident energy range of width 6.394 eV about the Fe L_3 resonance, including 30 grid points. The energy loss range is $E_{\text{loss}} \in [0.5 \text{ eV}, 2.2 \text{ eV}]$ measured on 151 grid points in total.

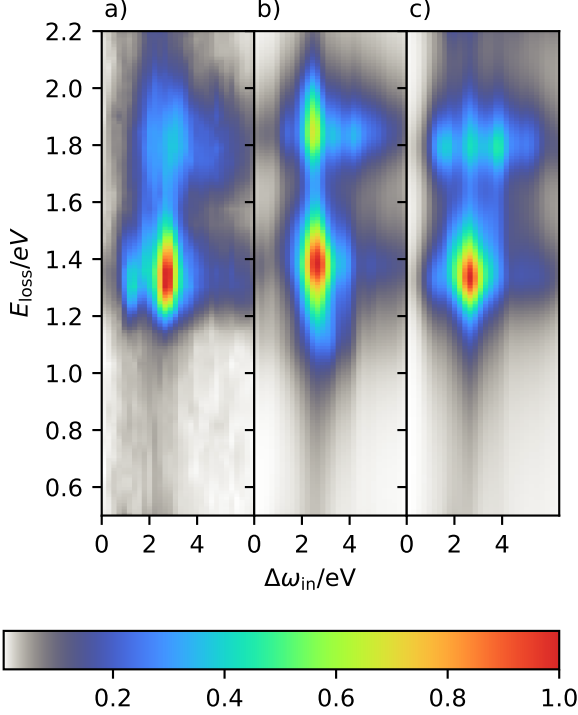


FIG. 5. Experimental and simulated spectra for Fe_2O_3 . (a) experimental data from Ref. [44], (b) result of GPR, (c) result after greedy fine-tuning (Powell’s method). L_1 max distances for panels (b)-(c) are $\chi^2_{L_1} = 158.8$ and 116.8, respectively.

2. GPR results

Figure 5 shows the experimental spectrum along with its simulated counterparts. Both the GPR and greedy fine-tuned fits reproduce many aspects of the experimental spectra. Parameter values corresponding to the greedy fit are reported in Tab. III. Ref. [44] attributed the two main branches of excitations in the $E_{\text{loss}} \sim 1 - 2$ eV range to dd excitations. They correspond to ${}^6A_{1g} \rightarrow {}^4T_{1g}$ (at 1.4 eV) and ${}^6A_{1g} \rightarrow {}^4T_{2g}$ (at 1.9 eV) transitions, flipping both spin ($S = 5/2 \rightarrow 3/2$) and orbital ($e_g \rightarrow t_{2g}$), respectively. This conclusion is largely consistent with the fit results, with the important remark that the fitted spectrum consists of multiple overlapping peaks, some of which feature ${}^2T_2(t_2)^5$ content.

D. $\text{Ca}_3\text{LiOsO}_6$

As our final test case, we consider Os L_3 -edge RIXS measurements on $\text{Ca}_3\text{LiOsO}_6$. The osmate material hosts Os^{5+} (d^3) ions that are surrounded by six O^{2-} ligands in an approximately octahedral OsO_6 environment. The octahedra themselves are arranged in a slightly distorted hexagonal $R\bar{3}c$ lattice but remaining very close to ideal Oh symmetry. $\text{Ca}_3\text{LiOsO}_6$ serves as a model $5d^3$

system in which the local OsO_6 octahedra are relatively isolated, allowing direct access to single ion physics. Previous RIXS measurements on this material have revealed a dramatic spin-orbit-induced splitting of the t_{2g} manifold [45]. As with the hematite case, no full d -shell atomic models have been reported for this compound.

1. Experimental setup

We use the experimental spectrum reported in Ref. [45] for this test case. The RIXS measurements were performed at the Advanced Photon Source (APS) at Sector 27 using the MERIX instrumentation. The experiment was performed at the Os L_3 edge using a powdered sample. The spectrum was recorded using a temperature of $T = 300$ K, scattering angles $\theta_{\text{in}} = 45^\circ$, $2\Theta = 90^\circ$, and an azimuth angle of $\varphi = 0^\circ$. Intensities were recorded in the incident energy range of width 8 eV about the Os L_3 resonance, including 5 grid points. The energy loss range is $E_{\text{loss}} \in [0.5 \text{ eV}, 5.0 \text{ eV}]$, including 151 grid points in total.

2. GPR results

Figure 6 shows the experimental RIXS spectra as well as simulated spectra. The overall structure of the data are well reproduced by the fitted models. The corresponding greedy-fitted parameters, along with their confidence intervals are included in Tab. IV. Ref. [45] reported four peaks present at $E_{\text{loss}} < 2$ eV that resonate at $\omega_{\text{in}} = 10.874$ keV, and a feature at $E_{\text{loss}} \approx 4.5$ eV that resonates at $\omega_{\text{in}} = 10.878$ keV. This behavior indicates that the features below 2 eV are intra- t_{2g} excitations, whereas the higher energy feature is a transition between the t_{2g} and e_g states.

Ref. [45] fixes the value of the crystal field to $10Dq = 4.5$ eV; our fitted $10Dq = 4.1$ eV is somewhat lower. We assign the dominant ground state character to an ${}^4A_{2g}$ state, in agreement with Ref. [45], although we find a more significant contribution of 2T_2 states mixed in Ref. [45] refers to the first two peaks as hybridized 2E_g and ${}^2T_{1g}$ states. The numerics (including simulations we conducted using the parameters in Ref. [45]), however, predict a more diverse hybridization of 2E_g and ${}^2T_{1g}$, ${}^4A_{2g}$ and 2T_2 . The second peak is a mixture of 2T_1 and 2E , in agreement with the prior work. The third peak is consistently evaluated to be an almost pure 2T_2 state. The fourth peak is predominantly 2T_2 , with some 2T_1 and 2E mixed in. This assignment is also shared between Ref. [45] and our fits.

V. DISCUSSION AND CONCLUSIONS

Our results demonstrate that it is possible to automate model Hamiltonian extraction from information-

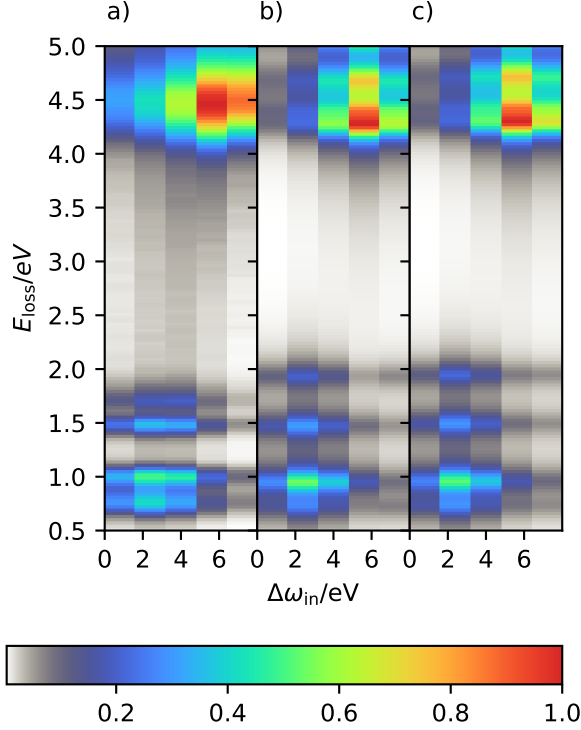


FIG. 6. Experimental and simulated spectra for $\text{Ca}_3\text{LiOsO}_6$. (a) experimental data from Ref. [45], (b) hand-fitted (reference), (c) result of GPR, result after greedy fine-tuning (Powell’s method). L_1 sum distances for panels (b)-(c) are $\chi^2_{L_1} = 0.313$ and 0.302 , respectively.

dense RIXS spectra of real materials, dramatically reducing the amount of effort and expertise required compared to traditional hand-tuned fitting.

A common challenge in model Hamiltonian extraction is the non-uniqueness of the inverse problem. Our work shows that multiple distinct parameter configurations can yield visually similar RIXS spectra. This problem is part of why hand-fitting RIXS spectra is challenging since it is difficult for humans to anticipate the number of different local minima of the distance function. Our automated analysis system enables a more comprehensive identification of these local minima. We further outline an automated means to use symmetry labels to more definitively identify distinct solutions.

This research points towards several promising directions for future work. One important issue concerns the choice of the distance function used to quantify similarity between the experimental and simulated spectra. Experimental data are often noisy, and simulated spectra can be affected by systematic model errors, making this comparison nontrivial. In this study, we employed pixel-wise metrics as a straightforward and transparent way to assess visual similarity. While effective in many cases, these metrics are sensitive to local fluctuations and may under represent high-level spectral features. A po-

TABLE IV. $\text{Ca}_3\text{LiOsO}_6$ parameters (in units of eV). The “interval 1.25” and “interval 1.5” refer to the range of values for each parameter that keep the distance function within $1.25\times$ or $1.5\times$ its minimum value, respectively, with all other parameters held fixed.

Parameter	Pt. est.	interval 1.25	interval 1.5
F_{dd}^2	2.78	[2.17, 3.37]	[1.53, 3.70]
F_{dp}^2	0.083	[0, 11.61]	[0, 19.17]
F_{dd}^4	1.37	[0.64, 2.25]	[0.01, 2.62]
G_{dp}^1	0.84	[n/a, 1.92]	N/A
G_{dp}^3	0.24	N/A	N/A
$10Dq$	4.15	[4.02, 4.29]	[3.91, 4.35]
$\Delta\omega_{in}$	-3.34	[-4.41, -2.00]	[-4.98, -1.38]
ζ_i	0.474	[0.393, 0.542]	[0.313, 0.581]
ζ_n	0.551	[0.206, 0.862]	N/A
ζ_c	N/A	N/A	N/A
Γ_n	3.0 (fixed)		

tentially promising extension of our work would be to incorporate learned perceptual similarity metrics, such as those based on deep neural networks, which have shown remarkable effectiveness in natural image comparison tasks [60]. Whether this approach would improve the quality of the workflow sufficient to justify the additional computational expense is an open question.

An obvious avenue to improve the method is to apply it to models of increasing sophistication such as Anderson impurity models or extended clusters. Such generalizations can provide a better approximation to the experimental spectrum over a wider range of energy loss at the cost of a larger number of floating parameters and a more expensive distance function. Further improvements to the active learning method (such as using deep Gaussian process models or refined distance functions) as well as numerical optimizations of the RIXS simulation together could represent suitable approaches to overcoming these difficulties. Simultaneously incorporating additional information to better constrain the problem is likely to be valuable. This could include first-principles calculations of tight-binding parameters and effective interactions.

In summary, this work establishes that a Bayesian optimization framework can automatically infer multi-orbital Hamiltonian parameters directly from information-dense RIXS spectra with an accuracy that rivals expert hand-fitting. Applied to four representative quantum materials— NiPS_3 , NiCl_2 , Fe_2O_3 , and $\text{Ca}_3\text{LiOsO}_6$ —the method not only reproduces published parameter sets but also supplies the first quantitative atomic-model parameters for Fe_2O_3 and $\text{Ca}_3\text{LiOsO}_6$. By rigorously mapping uncertainties, the approach offers a transparent route to evaluating model uniqueness and reliability. These results pave the way for extending inverse-problem automation to more elaborate cluster or impurity models and, ultimately, to high-throughput exploration of novel quantum materials with minimal human intervention.

CODE AND DATA AVAILABILITY

The code and data supporting this study will be deposited in a public online repository once the final version of the paper is accepted for publication. Until that time, both will be made available upon reasonable request.

ACKNOWLEDGMENTS

This work was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Award Number DE-SC0022311. We acknowledge Yilin Wang, Gilberto Fabbri, Gabi Kotliar, Max Rakitin, and Tom Hopkins for their contributions to EDRIXS. We thank Valentina Bisogni, Jiemin Li, and Johnny Pellicari for sharing data on Fe_2O_3 [44], Wei He for sharing data on NiPS_3 [13], Connor Occhialini and Riccardo Comin for sharing data on NiCl_2 [14], and Andy Christianson, and Stuart Calder for sharing data on $\text{Ca}_3\text{LiOsO}_6$ [45].

Appendix A: Detailed fitting information

1. NiPS_3

a. Fitting procedure

We used the sum-normalized L1 distance χ_{L1}^2 for this analysis. Figure 7 shows the decrease in the best probed distance function as a function of the GPR iterations for $N_{\text{run}} = 60$ independent runs. The variance of the best estimate after a fixed number of iterations is significant, motivating the use of multiple runs.

The best 73 points (with distances smaller than $1.15 \chi_{L1, \text{min}, \text{GPR}}^2$) are shown in the top panel of Fig. 8, as a function of F_{dd}^2 , F_{dd}^4 and $10Dq$. These points form

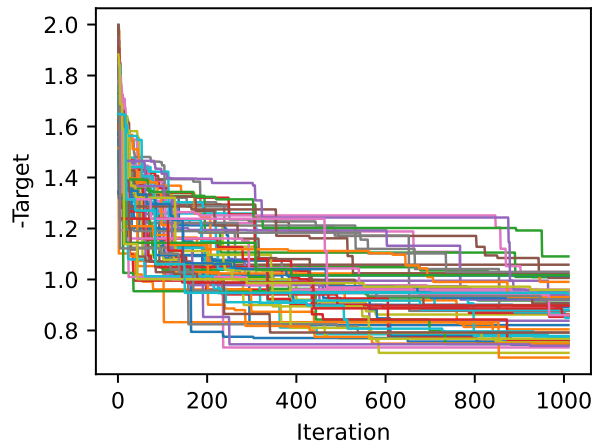


FIG. 7. NiPS_3 : Decrease of the sum normalized distance function for 60 runs of 1000 iterations

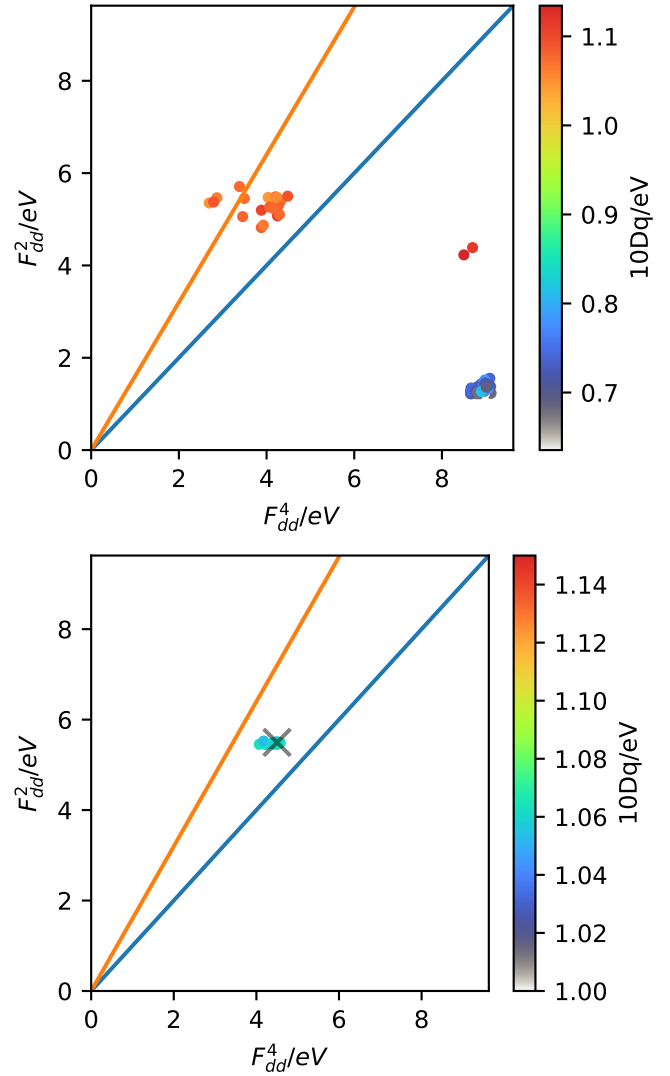


FIG. 8. Top: distribution of 18 GPR evaluations with $d \leq 1.15 \chi_{L1, \text{min}, \text{GPR}}^2$ for NiPS_3 . Bottom: results of subsequent greedy optimization starting from the 23 best GPR points. The chosen point is denoted by a gray \times . On both plots, the lines $F_{dd}^4 = F_{dd}^2$ and $F_{dd}^4 = 0.625 F_{dd}^2$ are drawn in blue and orange, respectively. The blue line is used to restrict candidate solutions, while the orange is shown as a guideline only.

three well-separated clusters, two of which are discarded due to unphysical F_{dd}^4/F_{dd}^2 ratios. After throwing out the unphysical point, we are left with $N_{\text{greedy}} = 23$ points, which are further refined with the greedy approach. The bottom panel of Fig. 8 shows the result of the greedy optimization starting from the best points of the GPR fit. All points collapse to the same valley of the distance function.

Figure 9 depicts the dependency of the distance functions on each floated parameter, keeping all the rest at their minimum values.

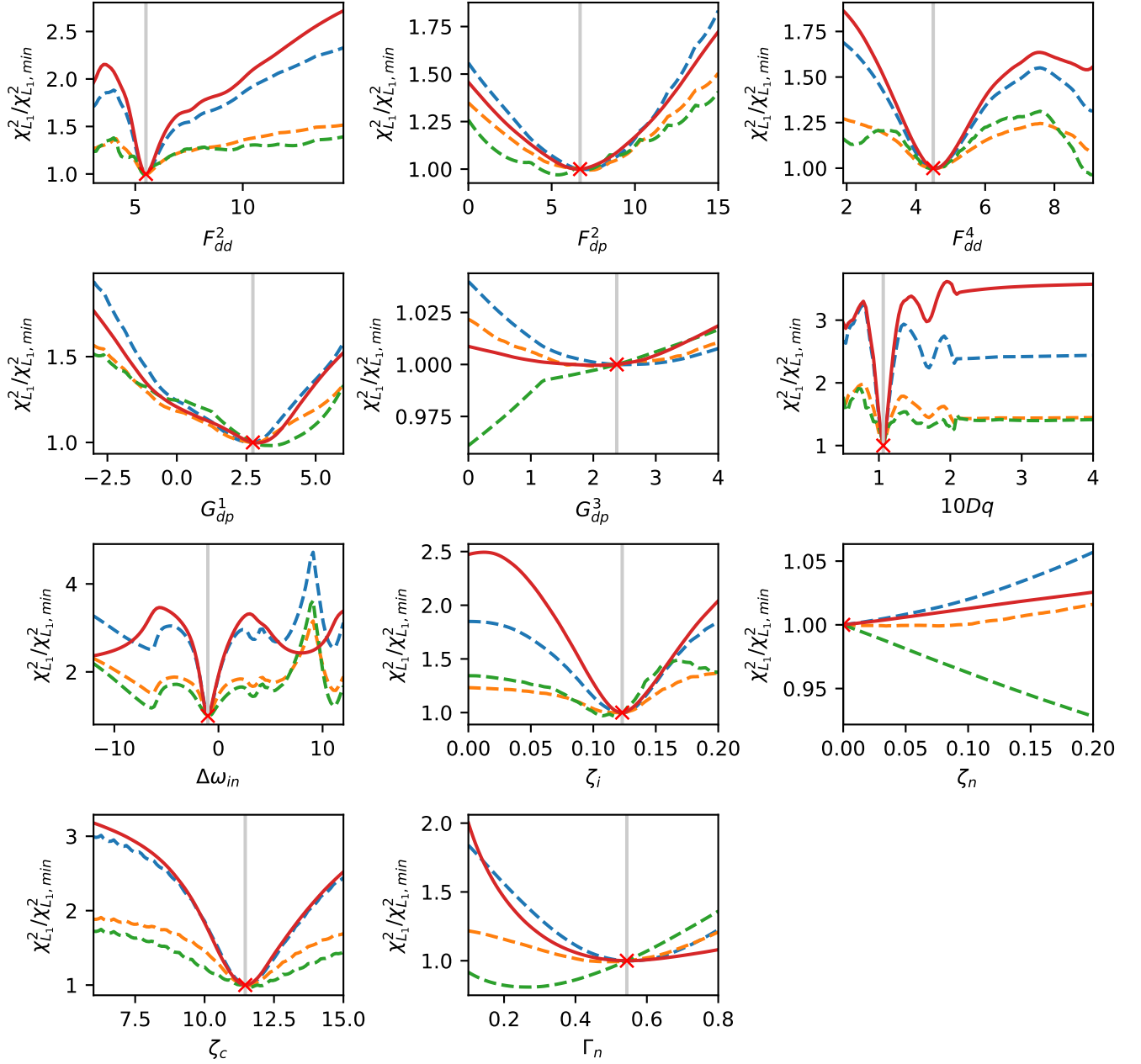


FIG. 9. NiPS₃: The behavior of various distance measures around the fine-tuned minimum. Solid red: L1 sum normalized, dashed blue: L1 maximum normalized, dashed orange: L2 sum normalized, dashed green: magnitude of gradient, maximum normalized. The distance functions are generally sensitive to the initial Slater parameters as well as $10Dq$ and the energy offset x_{offset} , while they are less sensitive to intermediate state parameters, especially ζ_n .

2. NiCl₂

a. Fitting procedure

We used the sum normalized L1 distance χ^2_{L1} for this analysis. Figure S1 of the SM [50] shows the decrease of the best probed distance function as function of the GPR iterations for $N_{\text{run}} = 60$ independent runs. The best 11 points (with distances smaller than $1.2\chi_{\text{min,GPR}}$)

are shown on the top panel of Fig. S2 (SM [50]), shown as function of F_{dd}^2 , F_{dd}^4 and $10Dq$. The bottom panel of Fig. S2 (SM [50]) exhibits the result of the greedy optimization starting from the outcomes of the GPR fit. Upon the greedy refinement, these points arrange into a single cluster.

Figure S3 of the SM [50] depicts the dependency of the distance functions on each one floated parameter, keeping all the rest at their minimum values.

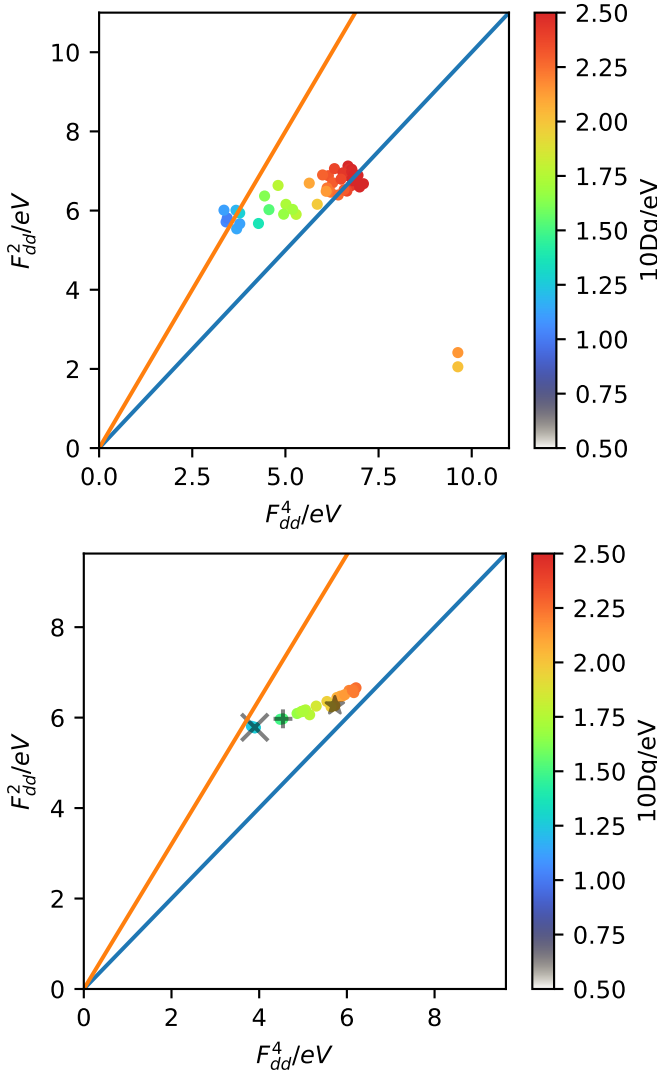


FIG. 10. Top: distribution of 50 GPR evaluations with $d \leq 1.2\chi_{L_1, \min, GPR}^2$ and $10Dq < 2.5$ for Fe_2O_3 . Bottom: results of subsequent greedy optimization starting from the 28 best GPR points. The overall best point is denoted by a gray X, while alternative solutions are marked with a plus and an asterisk. On both plots, the lines $F_{dd}^4 = F_{dd}^2$ and $F_{dd}^4 = 0.625F_{dd}^2$ are drawn in blue and orange, respectively. The blue line is used to restrict candidate solutions, while the orange is shown as a guideline only.

3. Fe_2O_3

a. Fitting procedure

We used the maximum normalized L1 distance $\chi_{L_1}^2$ for this analysis. The top panel of Fig. S4 (see SM [50]) shows the decrease of the best probed distance function as function of the GPR iterations for $N_{\text{run}} = 60$ independent runs.

The best 50 points (with distances smaller than $1.2\chi_{L_1, \min, GPR}^2$ and $10Dq < 2.5$) are shown on the top

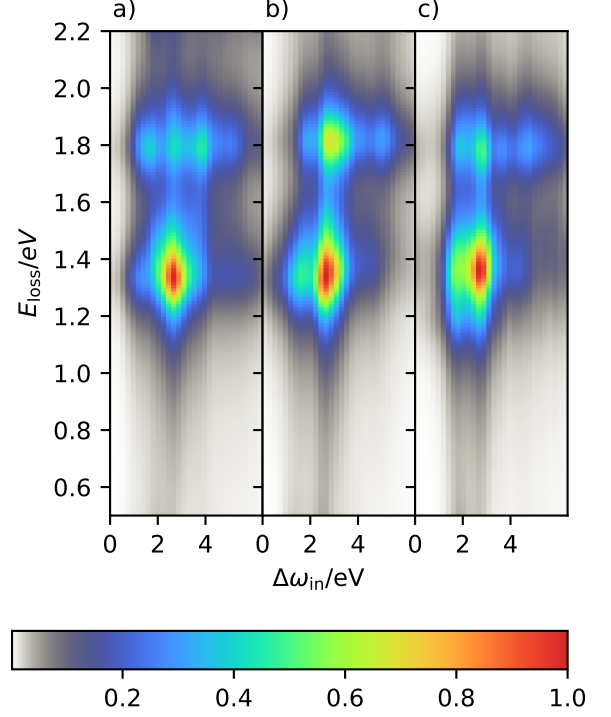


FIG. 11. Fe_2O_3 : Alternative results for the greedy fit. L1 max distances for panels (a)-(c) are $\chi_{L_1}^2 = 116.8, 129.6$ and 148.7 , respectively. The fits shown in panels a-c correspond to the parameter values indicated by the cross (×), plus (+), and star (★) in Fig. 10.

panel of Fig. 10. A subset of the candidate points have an unphysical F_{dd}^2/F_{dd}^4 ratio. We thus restrict our analysis to 28 points with $F_{dd}^2 \leq F_{dd}^4$ in the following. The bottom panel of Fig. S4 (SM [50]) depicts these remaining candidate points after the restriction.

The bottom panel of Fig. 10 exhibits the result of the greedy optimization starting from the outcomes of the GPR fit. Three separate but closely packed clusters can be distinguished on this plot. Figure 11 shows alternative fits corresponding to different minima of the greedy optimization.

Figure S5 in the SM [50] plots the dependency of the distance functions on each one floated parameter, keeping all the rest at their minimum values.

4. $\text{Ca}_3\text{LiOsO}_6$

a. Fitting procedure

We used the sum normalized L1 distance $\chi_{L_1}^2$ for this analysis. Figure S6 in the SM [50] shows the decrease of the best probed distance function as function of the GPR iterations for $N_{\text{run}} = 60$ independent runs. The best 24 points (with distances smaller than $1.06\chi_{L_1, \min, GPR}^2$) are

shown on the left panel of Fig. S7 (see SM [50]), shown as function of F_{dd}^2 , F_{dd}^4 and $10Dq$.

The right panel of Fig. S7 (SM [50]) exhibits the result of the greedy optimization starting from the outcomes of the GPR fit. The refined points organize into a single

cluster.

Figure S8 (SM [50]) depicts the dependency of the distance functions on each one floated parameter, keeping all the rest at their minimum values.

-
- [1] D. Basov, R. Averitt, and D. Hsieh, Towards properties on demand in quantum materials, *Nat. Mater.* **16**, 1077 (2017).
 - [2] Y. Tokura, M. Kawasaki, and N. Nagaosa, Emergent functions of quantum materials, *Nat. Phys.* **13**, 1056 (2017).
 - [3] F. Giustino, J. H. Lee, F. Trier, M. Bibes, S. M. Winter, R. Valentí, Y.-W. Son, L. Taillefer, C. Heil, A. I. Figueroa, B. Plaçais, Q. Wu, O. V. Yazyev, E. P. A. M. Bakkers, J. Nygård, P. Forn-Díaz, S. D. Franceschi, J. W. McIver, L. E. F. F. Torres, T. Low, A. Kumar, R. Galceran, S. O. Valenzuela, M. V. Costache, A. Manchon, E.-A. Kim, G. R. Schleder, A. Fazzio, and S. Roche, The 2021 quantum materials roadmap, *J. Phys. Mater.* **3**, 042006 (2021).
 - [4] M. Mitrano, S. Johnston, Y.-J. Kim, and M. P. M. Dean, Exploring quantum materials with resonant inelastic x-ray scattering, *Phys. Rev. X* **14**, 040501 (2024).
 - [5] P. W. Anderson, *Concepts in Solids* (World Scientific, 1997).
 - [6] E. Gull, O. Parcollet, and A. J. Millis, Superconductivity and the pseudogap in the two-dimensional Hubbard model, *Phys. Rev. Lett.* **110**, 216405 (2013).
 - [7] H.-C. Jiang and T. P. Devereaux, Superconductivity in the doped Hubbard model and its interplay with next-nearest hopping t' , *Science* **365**, 1424 (2019).
 - [8] M. Qin, C.-M. Chung, H. Shi, E. Vitali, C. Hubig, U. Schollwöck, S. R. White, and S. Zhang (Simons Collaboration on the Many-Electron Problem), Absence of superconductivity in the pure two-dimensional Hubbard model, *Phys. Rev. X* **10**, 031016 (2020).
 - [9] Z. Chen, Y. Wang, S. N. Rebec, T. Jia, M. Hashimoto, D. Lu, B. Moritz, R. G. Moore, T. P. Devereaux, and Z.-X. Shen, Anomalous strong near-neighbor attraction in doped 1D cuprate chains, *Science* **373**, 1235 (2021).
 - [10] H. Padma, J. Thomas, S. F. R. TenHuisen, W. He, Z. Guan, J. Li, B. Lee, Y. Wang, S. H. Lee, Z. Mao, H. Jang, V. Bisogni, J. Pellicciari, M. P. M. Dean, S. Johnston, and M. Mitrano, Beyond-hubbard pairing in a cuprate ladder, *Phys. Rev. X* **15**, 021049 (2025).
 - [11] A. Scheie, P. Laurell, J. Thomas, V. Sharma, A. I. Kolesnikov, G. E. Granroth, Q. Zhang, B. Lake, M. M. Jr., R. I. Bewley, R. S. Eccleston, J. Akimitsu, E. Dagotto, C. D. Batista, G. Alvarez, S. Johnston, and D. A. Tennant, Cooper-pair localization in the magnetic dynamics of a cuprate ladder, *arXiv:2501.10296* (2025).
 - [12] S. Kang, K. Kim, B. H. Kim, J. Kim, K. I. Sim, J.-U. Lee, S. Lee, K. Park, S. Yun, T. Kim, A. Nag, A. Walters, M. Garcia-Fernandez, J. Li, L. Chapon, K.-J. Zhou, Y.-W. Son, J. H. Kim, H. Cheong, and J.-G. Park, Coherent many-body exciton in van der Waals antiferromagnet NiPS₃, *Nature* **583**, 785 (2020).
 - [13] W. He, Y. Shen, K. Wohlfeld, J. Sears, J. Li, J. Pellicciari, M. Walicki, S. Johnston, E. Baldini, V. Bisogni, M. Mitrano, and M. P. M. Dean, Magnetically propagating Hund's exciton in van der Waals antiferromagnet NiPS₃, *Nature Communications* **15** (2024).
 - [14] C. A. Occhialini, Y. Tseng, H. Elnaggar, Q. Song, M. Blei, S. A. Tongay, V. Bisogni, F. M. F. de Groot, J. Pellicciari, and R. Comin, Nature of excitons and their ligand-mediated delocalization in nickel dihalide charge-transfer insulators, *Phys. Rev. X* **14**, 031007 (2024).
 - [15] I. J. Hamad, C. S. Helman, L. O. Manuel, A. E. Feiguin, and A. A. Aligia, Singlet polaron theory of low-energy optical excitations in NiPS₃, *Phys. Rev. Lett.* **133**, 146502 (2024).
 - [16] V. Sharma, Z. Wang, and C. D. Batista, Machine learning assisted derivation of minimal low-energy models for metallic magnets, *npj Computational Materials* **9**, 192 (2023).
 - [17] J. Carrasquilla, Machine learning for quantum matter, *Advances in Physics: X* **5**, 1797528 (2020).
 - [18] E. Bedolla, L. C. Padierna, and R. Castañeda-Priego, Machine learning for condensed matter physics, *Journal of Physics: Condensed Matter* **33**, 053001 (2020).
 - [19] M. R. Carbone, When not to use machine learning: A perspective on potential and limitations, *MRS Bulletin* **47**, 968–974 (2022).
 - [20] S. Johnston, E. Khatami, and R. Scalettar, A perspective on machine learning and data science for strongly correlated electron problems, *Carbon Trends* **9**, 100231 (2022).
 - [21] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, Machine learning and the physical sciences, *Rev. Mod. Phys.* **91**, 045002 (2019).
 - [22] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, Physics-informed machine learning, *Nature Reviews Physics* **3**, 422 (2021).
 - [23] Z. Chen, N. Andrejevic, N. C. Drucker, T. Nguyen, R. P. Xian, T. Smidt, Y. Wang, R. Ernstorfer, D. A. Tennant, M. Chan, and M. Li, Machine learning on neutron and x-ray scattering and spectroscopies, *Chemical Physics Reviews* **2**, 031301 (2021).
 - [24] M. R. Carbone, S. Yoo, M. Topsakal, and D. Lu, Classification of local chemical environments from x-ray absorption spectra using supervised machine learning, *Phys. Rev. Mater.* **3**, 033604 (2019).
 - [25] M. R. Carbone, M. Topsakal, D. Lu, and S. Yoo, Machine-learning x-ray absorption spectra to quantitative accuracy, *Phys. Rev. Lett.* **124**, 156401 (2020).
 - [26] S. B. Torrisi, M. R. Carbone, B. A. Rohr, J. H. Montoya, Y. Ha, J. Yano, S. K. Suram, and L. Hung, Random forest machine learning models for interpretable x-ray absorption near-edge structure spectrum-property relationships, *npj Comput. Mater.* **6**, 1 (2020).
 - [27] C. Miles, M. R. Carbone, E. J. Sturm, D. Lu, A. Weichselbaum, K. Barros, and R. M. Konik, Machine learn-

- ing of Kondo physics using variational autoencoders and symbolic regression, *Phys. Rev. B* **104**, 235111 (2021).
- [28] E. J. Sturm, M. R. Carbone, D. Lu, A. Weichselbaum, and R. M. Konik, Predicting impurity spectral functions using machine learning, *Phys. Rev. B* **103**, 245118 (2021).
- [29] H. Kwon, W. Sun, T. Hsu, W. Jeong, F. Aydin, S. Sharma, F. Meng, M. R. Carbone, X. Chen, D. Lu, *et al.*, Harnessing neural networks for elucidating x-ray absorption structure–spectrum relationships in amorphous carbon, *The Journal of Physical Chemistry C* **127**, 16473 (2023).
- [30] H. Kwon, T. Hsu, W. Sun, W. Jeong, F. Aydin, J. Chapman, X. Chen, V. Lordi, M. R. Carbone, D. Lu, F. Zhou, and T. Anh Pham, Spectroscopy-guided discovery of three-dimensional structures of disordered materials with diffusion models, *Machine Learning: Science and Technology* **5**, 045037 (2024).
- [31] A. Ghose, M. Segal, F. Meng, Z. Liang, M. S. Hybertsen, X. Qu, E. Stavitski, S. Yoo, D. Lu, and M. R. Carbone, Uncertainty-aware predictions of molecular x-ray absorption spectra using neural network ensembles, *Phys. Rev. Res.* **5**, 013180 (2023).
- [32] C. D. Rankine, M. M. Madkhali, and T. J. Penfold, A deep neural network for the rapid prediction of x-ray absorption spectra, *The Journal of Physical Chemistry A* **124**, 4263 (2020).
- [33] R. F. L. Evans, W. J. Fan, P. Chureemart, T. A. Ostler, M. O. A. Ellis, and R. W. Chantrell, Atomistic spin model simulations of magnetic nanomaterials, *Journal of Physics: Condensed Matter* **26**, 103202 (2014).
- [34] S. Toth and B. Lake, Linear spin wave theory for single-Q incommensurate magnetic structures, *Journal of Physics: Condensed Matter* **27**, 166002 (2015).
- [35] A. Nocera and G. Alvarez, Spectral functions with the density matrix renormalization group: Krylov-space approach for correction vectors, *Phys. Rev. E* **94**, 053308 (2016).
- [36] D. Dahlbom, H. Zhang, C. Miles, S. Quinn, A. Niraola, B. Thipe, M. Wilson, S. Matin, H. Mankad, S. Hahn, D. Pajerowski, S. Johnston, Z. Wang, H. Lane, Y. W. Li, X. Bai, M. Mourigal, C. D. Batista, and K. Barros, Sunny.jl: A julia package for spin dynamics, *arXiv:2501.13095* (2025).
- [37] A. M. Samarakoon, K. Barros, Y. W. Li, M. Eisenbach, Q. Zhang, F. Ye, V. Sharma, Z. L. Dun, H. Zhou, S. A. Grigera, C. D. Batista, and D. A. Tennant, Machine-learning-assisted insight into spin ice $\text{Dy}_2\text{Ti}_2\text{O}_7$, *Nature Communications* **11**, 892 (2020).
- [38] A. M. Samarakoon and D. A. Tennant, Machine learning for magnetic phase diagrams and inverse scattering problems, *Journal of Physics: Condensed Matter* **34**, 044002 (2021).
- [39] A. Samarakoon, D. A. Tennant, F. Ye, Q. Zhang, and S. A. Grigera, Integration of machine learning with neutron scattering for the hamiltonian tuning of spin ice under pressure, *Communications Materials* **3**, 84 (2022).
- [40] Y. Cao, D. G. Mazzone, D. Meyers, J. P. Hill, X. Liu, S. Wall, and M. P. M. Dean, Ultrafast dynamics of spin and orbital correlations in quantum materials: An energy- and momentum-resolved perspective, *Philos. Trans. R. Soc. A* **377**, 20170480 (2019).
- [41] M. Mitrano and Y. Wang, Probing light-driven quantum materials with ultrafast resonant inelastic x-ray scattering, *Comm. Phys.* **3**, 184 (2020).
- [42] Y. L. Wang, G. Fabbris, M. P. M. Dean, and G. Kotliar, EDRIXS: An open source toolkit for simulating spectra of resonant inelastic x-ray scattering, *Computer Physics Communications* **243**, 151 (2019).
- [43] P. Zimmermann, R. J. Green, M. W. Haverkort, and F. M. F. de Groot, *Quanta4RIXS*: a program for crystal field multiplet calculations of RIXS and RIXS–MCD spectra using *Quanta*, *Journal of Synchrotron Radiation* **25**, 899 (2018).
- [44] J. Li, Y. Gu, Y. Takahashi, K. Higashi, T. Kim, Y. Cheng, F. Yang, J. Kuneš, J. Pelliciari, A. Hariki, and V. Bisogni, Single- and Multimagnon Dynamics in Antiferromagnetic α - Fe_2O_3 Thin Films, *Physical Review X* **13**, 011012 (2023).
- [45] A. E. Taylor, S. Calder, R. Morrow, H. L. Feng, M. Upton, M. D. Lumsden, K. Yamaura, P. M. Woodward, A. D. Christianson, and A. D. Christianson, Spin-orbit coupling controlled $J = 3/2$ electronic ground state in $5d^3$ oxides., *Physical review letters* **118** **20**, 207202 (2016).
- [46] Y. Wang, G. Fabbris, M. Dean, and G. Kotliar, EDRIXS: An open source toolkit for simulating spectra of resonant inelastic x-ray scattering, *Computer Physics Communications* **243**, 151 (2019).
- [47] EDRIXS website, <https://github.com/NSLS-II/edrixx> (2025), accessed: 2025-06-01.
- [48] L. J. P. Ament, M. van Veenendaal, T. P. Devereaux, J. P. Hill, and J. van den Brink, Resonant inelastic x-ray scattering studies of elementary excitations, *Reviews of Modern Physics* **83**, 705 (2011).
- [49] F. de Groot and A. Kotani, *Core Level Spectroscopy of Solids* (CRC Press, 2008).
- [50] See the online supplementary materials, available at: [URLsuppliedbythepublisher](https://www.nature.com/articles/s41534-025-00000-0).
- [51] P. I. Frazier, A tutorial on Bayesian optimization, *arXiv:1807.02811* (2018).
- [52] J. Wang, An Intuitive Tutorial to Gaussian Process Regression, *Computing in Science and Engineering* **25**, 4 (2023).
- [53] C. E. Rasmussen and C. K. I. Williams, *Gaussian process for machine learning* (The MIT Press, 2006).
- [54] E. Kaufmann, O. Cappe, and A. Garivier, On bayesian upper confidence bounds for bandit problems, in *Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics*, Proceedings of Machine Learning Research, Vol. 22, edited by N. D. Lawrence and M. Girolami (PMLR, La Palma, Canary Islands, 2012) pp. 592–600.
- [55] C. Cartis, J. Fiala, B. Marteau, and L. Roberts, Improving the flexibility and robustness of model-based derivative-free optimization solvers, *ACM Trans. Math. Softw.* **45**, 10.1145/3338517 (2019).
- [56] A. R. Wildes, J. R. Stewart, M. D. Le, R. A. Ewings, K. C. Rule, G. Deng, and K. Anand, Magnetic dynamics of NiPS_3 , *Phys. Rev. B* **106**, 174422 (2022).
- [57] A. R. Wildes, V. Simonet, E. Ressouche, G. J. McIntyre, M. Avdeev, E. Suard, S. A. J. Kimber, D. Lançon, G. Pepe, B. Moubaraki, and T. J. Hicks, Magnetic structure of the quasi-two-dimensional antiferromagnet NiPS_3 , *Phys. Rev. B* **92**, 224408 (2015).
- [58] K. Kim, S. Y. Lim, J.-U. Lee, S. Lee, T. Y. Kim, K. Park, G. S. Jeon, C.-H. Park, J.-G. Park, and H. Cheong, Suppression of magnetic ordering in XXZ -type antiferromagnetic monolayer NiPS_3 , *Nature Communications* **10**,

- 345 (2019).
- [59] S. Y. Kim, T. Y. Kim, L. J. Sandilands, S. Sinn, M.-C. Lee, J. Son, S. Lee, K.-Y. Choi, W. Kim, B.-G. Park, C. Jeon, H.-D. Kim, C.-H. Park, J.-G. Park, S. J. Moon, and T. W. Noh, Charge-spin correlation in van der Waals antiferromagnet NiPS₃, *Phys. Rev. Lett.* **120**, 136402 (2018).
- [60] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, The unreasonable effectiveness of deep features as a perceptual metric, in *CVPR* (2018).

Supplementary Material for “Hamiltonian parameter inference from RIXS spectra with active learning”

Marton K. Lajer^{1,*}, Xin Dai², Kipton Barros³,
Matthew R. Carbone², S. Johnston^{4,5,†} and M. P. M. Dean^{1,4,‡}

¹Condensed Matter Physics and Materials Science Department, Brookhaven National Laboratory, Upton, New York 11973, USA

²Computing and Data Sciences Directorate, Brookhaven National Laboratory, Upton, New York 11973, USA

³Theoretical Division and CNLS, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

⁴Department of Physics and Astronomy, The University of Tennessee, Knoxville, Tennessee 37996, USA

⁵Institute for Advanced Materials and Manufacturing, University of Tennessee, Knoxville, Tennessee 37996, USA

(Dated: August 4, 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 be obtained from the eigenvalues of A :

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

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

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}_λ of characters. The multiplicity of each irrep in the subspace is given by the expression

$$\mu_\lambda = \mathcal{X}_\lambda T^{-1}. \quad (1)$$

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 $|\phi_{q_i;\lambda;r}\rangle$

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

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

$$\sum_{q_i;\lambda} w_{\psi,q_i;\lambda} = 1. \quad (3)$$

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])

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

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

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

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 :

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

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

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

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 $O \cong S_4$

	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

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

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

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

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_α} onto the irrep $D_\alpha^{(k)}$ is

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

What is the largest amplitude of v_{m_α} 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

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

We would like to calculate

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

Introducing $V_{lj} = v_{m_\alpha,j}^{(l)}$, we then obtain

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

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

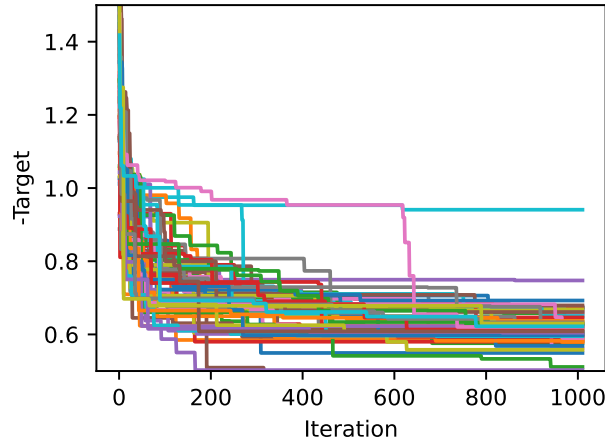


FIG. S1. NiCl_2 : Decrease of the sum normalized L_1 distance function for 60 runs of 1000 iterations

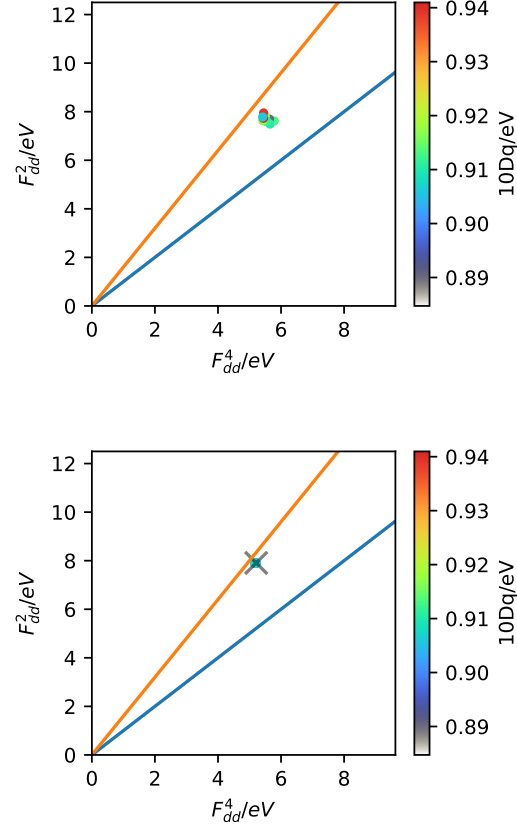


FIG. S2. Top: distribution of 11 GPR evaluations with $d \leq 1.2\chi_{\text{min,GPR}}^2$ for NiCl_2 . Bottom: results of subsequent greedy optimization starting from the 11 best GPR points. The chosen point is denoted by a gray cross (×).

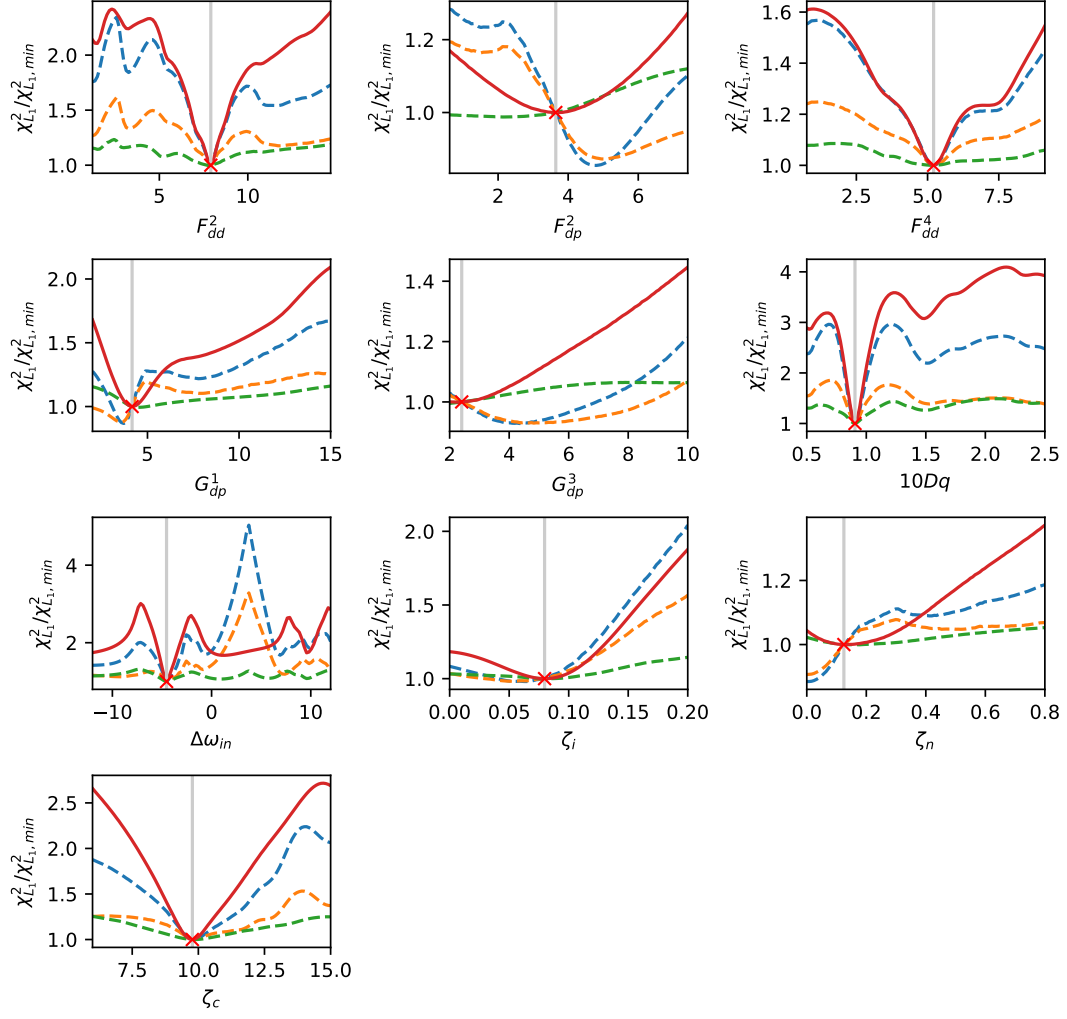


FIG. S3. NiCl_2 : The behavior of various distance measures around the fine-tuned minimum. Solid red: L1 sum normalized, dashed blue: L1 maximum normalized, dashed orange: L2 sum normalized, dashed green: magnitude of gradient, maximum normalized. The fit is highly sensitive to initial Slater and crystal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

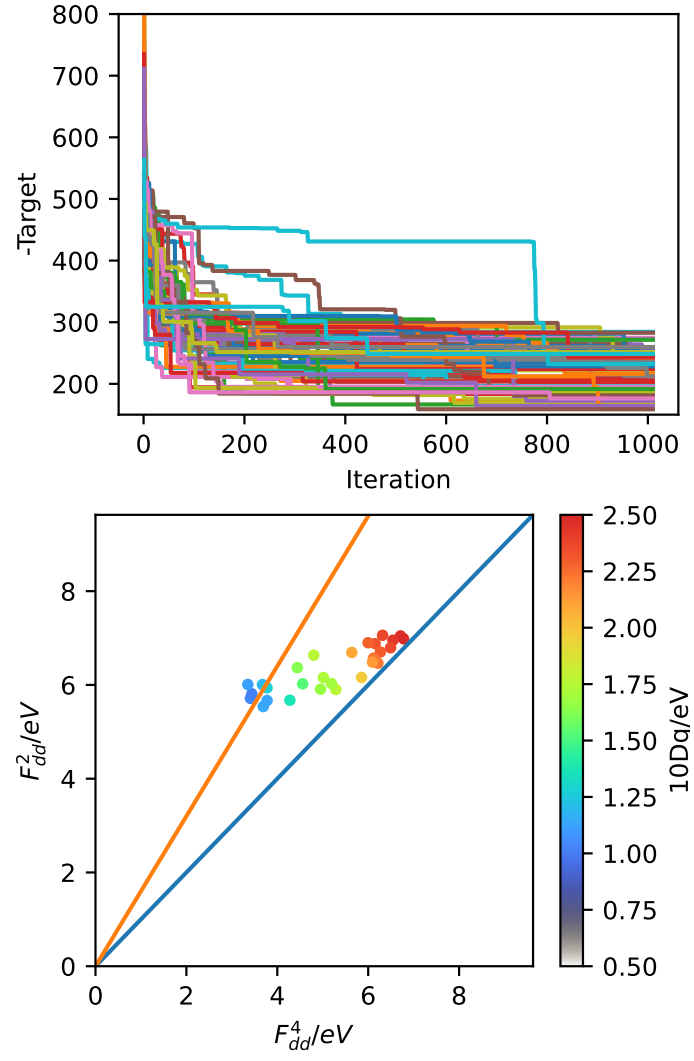
B. Fe_2O_3 

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.

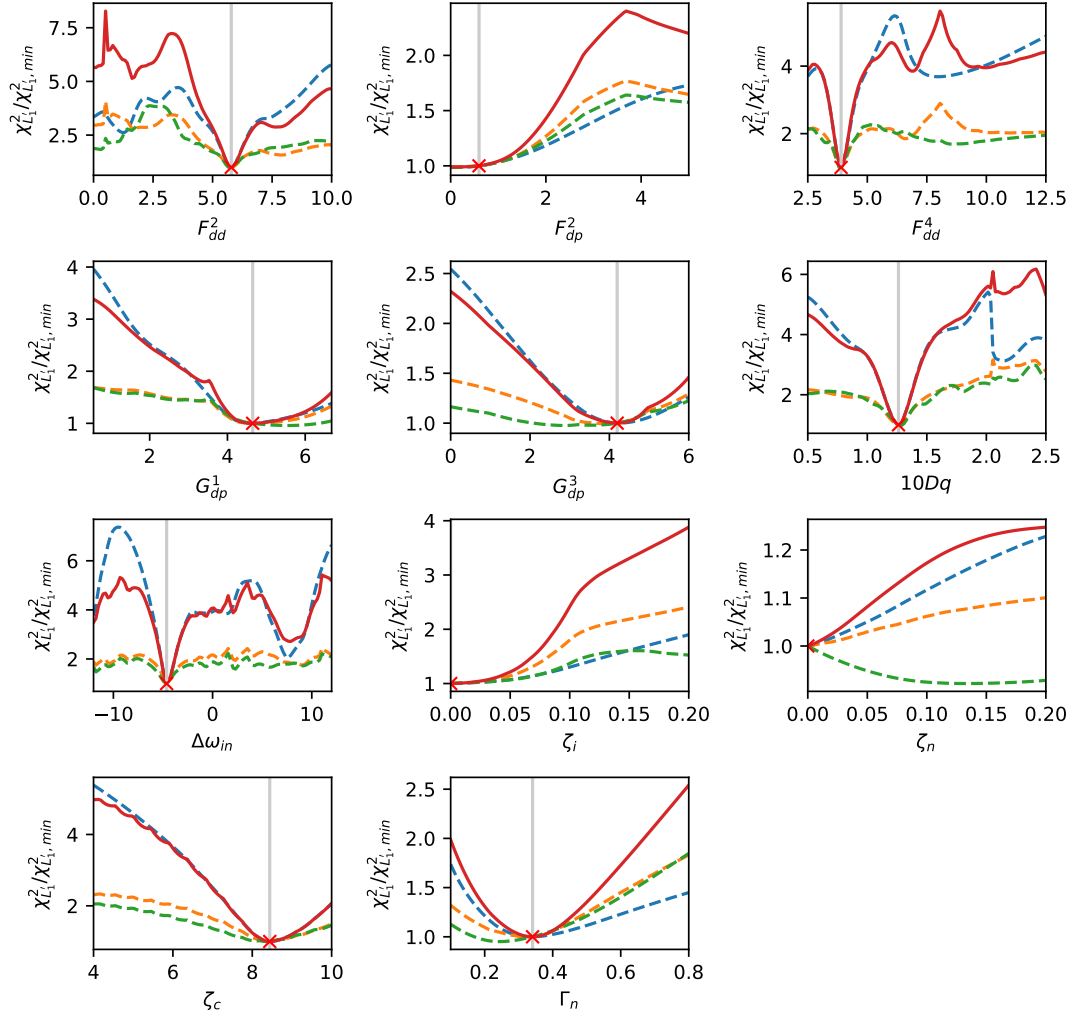
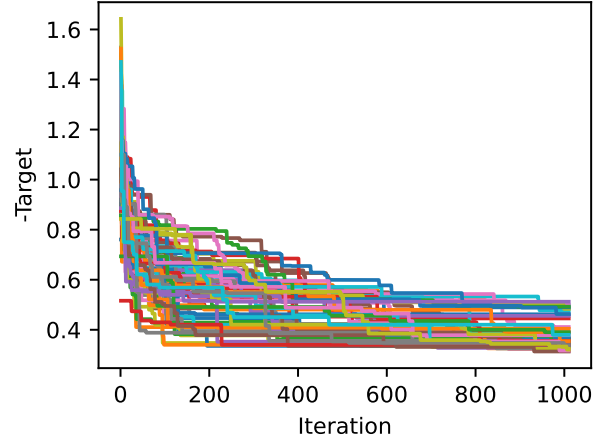
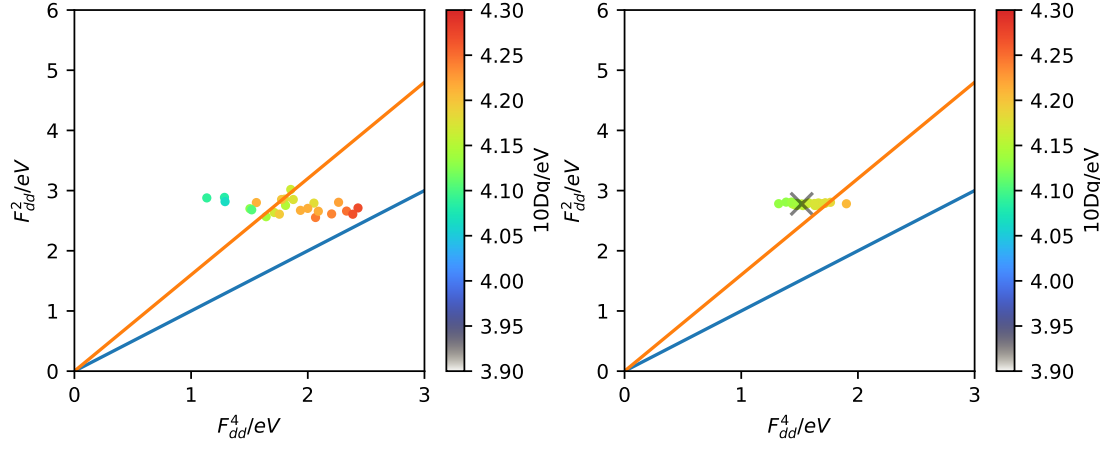


FIG. S5. Fe_2O_3 : The behavior of various distance measures around the fine-tuned minimum. Solid red: L1 *maximum* normalized, dashed blue: L1 sum normalized, dashed orange: L2 sum normalized, dashed green: magnitude of gradient, maximum normalized. The fit is highly sensitive to initial Slater and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

C. $\text{Ca}_3\text{LiOsO}_6$ FIG. S6. $\text{Ca}_3\text{LiOsO}_6$: Decrease of the sum normalized L1 distance function for 60 runs of 1000 iterationsFIG. S7. Left: distribution of 24 GPR evaluations with $d \leq 1.06\chi_{\min, \text{GPR}}^2$ for $\text{Ca}_3\text{LiOsO}_6$. Right: results of subsequent greedy optimization starting from the 24 best GPR points. The final accepted point is denoted by a gray cross (\times).

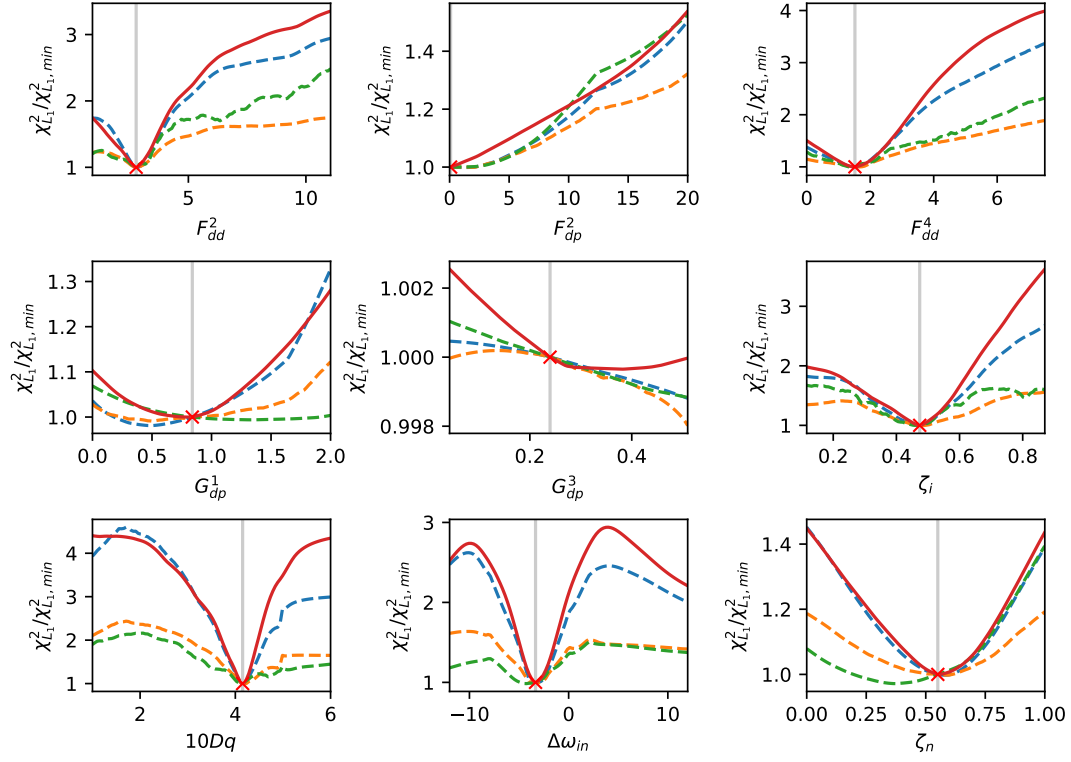


FIG. S8. $\text{Ca}_3\text{LiOsO}_6$: The behavior of various distance measures around the fine-tuned minimum. Solid red: L1 sum normalized, dashed blue: L1 maximum normalized, dashed orange: L2 sum normalized, dashed green: magnitude of gradient, maximum normalized. The fit is highly sensitive to initial 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).