

Determination of ligand field parameters of single-molecule magnets from magnetic susceptibility using deep learning based inverse models

Zayan A. Ali , Julius Mutschler , and Oliver Waldmann ^{*}
 Physikalisches Institut, *Universität Freiburg*, D-79104 Freiburg, Germany



(Received 28 April 2025; revised 25 June 2025; accepted 14 July 2025; published 4 August 2025)

Lanthanide-based single-molecule magnets (SMMs) have garnered extensive attention in recent decades for their potential applications in magnetic memory and quantum computing. In ligand field theory, the magnetic properties of a lanthanide ion may have to be characterized by up to 27 ligand field parameters. However, common experimental data such as temperature-dependent susceptibility curves measured on powder samples typically lack distinctive features. This leads to a massive over-parametrization, where multiple parameter sets can equally describe the data. In this work the issue is tackled by applying a deep learning architecture which combines a variational autoencoder (VAE) with an invertible neural network (INN). This VAE-INN architecture is shown to be capable of determining multiple ligand field parameter sets by extracting hidden parameters from the magnetic data and mapping them to the space of ligand field parameters. The approach is tested on simulated and experimental susceptibility data in magnetic models involving second-order ligand field parameters and is gradually extended to higher-order parameters until it is found limited by dimensionality issues. Compared with traditional least-squares fitting methods, the proposed approach exhibits enhanced generalization, robustness, and convergence. The present investigation of the VAE-INN architecture can be hoped to pave the way for further exploration of higher-order ligand field parameters.

DOI: [10.1103/7tfd-22jr](https://doi.org/10.1103/7tfd-22jr)

I. INTRODUCTION

Machine-learning techniques are rapidly establishing themselves as powerful tools in the fields of material science and condensed-matter physics, offering novel approaches to complex problems that are often challenging or even intractable by traditional means. In recent years, they have been successfully applied to quantum many-body systems [1–3], integrated with inelastic neutron scattering (INS) spectra [4,5], used to predict electronic and magnetic properties from first principles [6,7] and relating parameters to experimental observables [8–13]. In this work, a deep learning architecture is applied to a novel use case, the determination of magnetic parameters from temperature-dependent magnetic susceptibility curves of lanthanide-based single molecule magnets (SMMs) [14,15].

SMMs and in particular the lanthanide-based SMMs have attracted enormous attention due to their magnetic bistability [16–21] at the molecular level suggesting applications in data storage [22,23], spintronics [24], and quantum computing [25–27]. The magnetic properties of lanthanide ions are particularly sensitive to the splittings in their ground J multiplet arising from the $4f$ electron shell due to the ligand environment (J is the total angular-momentum quantum

number). These are commonly described by the ligand field Hamiltonian

$$\hat{H}_{\text{LF}} = \sum_{kq} B_{kq} \hat{O}_{kq}, \quad (1)$$

where $q = -k, \dots, k$, $0 < k \leq 2J$ and k is even (the $k = 0$ contribution can be omitted), $\hat{O}_{kq}$ are the Stevens operators, and B_{kq} are the ligand field parameters [28,29]. A widely employed approximation consists of considering only contributions of order $k \leq 6$, which results in a description of the ligand field effects with a set of 27 parameters [29]. Measurements of the magnetic susceptibility $\chi(T)$, which describes the magnetic response upon application of (small) magnetic field as a function of temperature T and which is a most common method for characterizing the magnetism in lanthanide-based complexes, are frequently performed on powdered samples, which produce relatively featureless data. Least-squares fitting algorithms such as Levenberg-Marquardt (LM) [30,31] can then fit only a few ligand field parameters using the information contained in such experimental data, and, furthermore, the selection of the parameters in the fitting model is neither unique nor known *a priori*. Symmetry arguments help to reduce the number of parameters, but this is less effective in the low-symmetry ligand environment often encountered in lanthanide-based SMMs [32–34]. These factors make the process of determining ligand field parameters from such experimental data a challenging task, even when more extensive datasets are available, because it requires the modeling of a function representative of a “one-to-many” mapping.

Ab initio methods like complete active space self-consistent field (CASSCF) approaches [35], have yielded valuable

^{*}Contact author: oliver.waldmann@physik.uni-freiburg.de

insights into the magnetic anisotropy in lanthanide-based SMMs [36–38]. Although these techniques are powerful, they are also complex and the accuracy of the theoretical predictions can be limited on account of necessary approximations [36,39–42]. A number of phenomenological techniques have been explored towards obtaining more manageable parametrizations of lanthanide SMMs [29]. These involve more intuitive approaches such as the point charge model (PCM) [43] or the angular overlap model (AOM), [44] and the many variations thereof [34,45–50]. Common to all these approaches is the search for a simpler parametrization of the ligand field, mostly through leveraging the known structure of the ligand environment, as obtained from x-ray crystal structure analysis. In this spirit, data-driven and machine-learning methods have shown promising successes [51–53]. For instance, Ref. [51] uses a machine-learning approach to predict the magnetic anisotropy from the chemical structure of single-ion magnets. Similarly, Ref. [54] uses a Gaussian-moment neural network to model the magnetic anisotropy tensor as well as the interatomic potentials of transition metal compounds. In these cases, the models are trained on data obtained from *ab initio* calculations to predict magnetic properties based on molecular parameters. While such methods show promising results, their extension to lanthanide-based SMMs remains limited due to difficult parametrization of higher-order ($k \geq 4$) contributions to the ligand field.

In this work, a machine learning approach is investigated which obtains a parametrization of the ligand field based not on ligand environment but simply on the statistical variance present in observable magnetic susceptibility data. To this end, in a first task it captures the essential variations in simulated $\chi(T)$ curves for a single-ion model in the form of “hidden parameters.” Such a hidden parametrization could be obtained, for example, through principle component analysis, but more powerful nonlinear techniques are available in the form of autoencoders [55]. This step can be regarded as an automated highly efficient parameter reduction step, which extracts a small number of “essential” parameters for modeling the magnetic susceptibility curves instead of the potentially many ligand field parameters. Since the hidden parameters are based on the statistical variance in the dataset, the resulting parametrization is not physically or chemically motivated. The second task is to relate the obtained hidden parametrization to the ligand field parameters B_{kq} . The over-parametrized nature of the problem means that possibly many different ligand field parameter sets will map to a single hidden parameter set. The learning of this relation will require the modeling of a surjective mapping, also known as an inverse problem or a one-to-many problem. The proposed deep learning solution for parametrizing the ligand field of SMMs using magnetic susceptibility data consists of combining a variational autoencoder (VAE) [55], which obtains the hidden statistical parametrization, with an invertible neural network (INN) [56], which models the inverse one-to-many mapping towards the ligand field parameters. A schematic of the developed VAE-INN architecture is presented in Fig. 1. During the training stage, the VAE encodes simulated powder susceptibility curves to a low-dimensional continuous latent space while maintaining reconstruction accuracy. The VAE’s latent variables function as the hidden parametrization of the

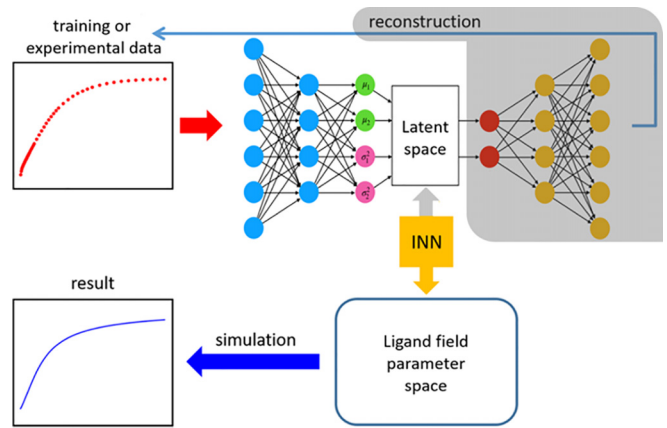


FIG. 1. Scheme of the VAE-INN architecture. The VAE encodes experimental temperature dependent magnetic susceptibility data (red scatter plot) to a unique point in the latent space (space of hidden parameters) for which the INN predicts all associated B_{kq} ligand field parameter sets. The parameter predictions are then used to simulate susceptibility curves (dark blue) which can be compared with the original input curve. During testing and for application, the components shaded in gray are not required.

susceptibility data. An INN is subsequently trained bidirectionally between the VAE latent space and the ligand field parameter space. For application, a susceptibility curve is supplied to the VAE, which encodes the curve to a unique point in the latent space for which the INN then predicts all associated ligand field parameter sets. For each of these parameter-set predictions, susceptibility curves can then be simulated and compared with the supplied input curve.

To study its capabilities, the trained VAE-INN model will be applied to simulated “experimental” data as well as to the experimental susceptibility curve of the “butterfly” molecule $[\text{Al}_2\text{Er}_2(\mu_3\text{-OH})_2(\text{pmide})_2(\text{p-Me-PhCO}_2)_6] \cdot 2\text{MeCN}$ ($\text{pmideH}_2 = \text{N-(2-pyridylmethyl)-iminodiethanol}$), or Al_2Er_2 in short [Fig. 2(a)] [57]. The simulated experimental data allows for a systematic investigation of the VAE-INN’s performance; the Al_2Er_2 $\chi(T)$ data showcases a practical application. Magnetic models of increasing complexity are considered, starting with models which include only the second-order B_{2q} ligand field parameters and later extending to models which include the effect of B_{4q} parameters. Interestingly, apart from being capable of predicting multiple ligand field parameter sets for a given susceptibility curve, it is found that the VAE-INN model is also easier to initialize and to make robust against various types of experimental errors compared with conventional least-squares fitting routines such as LM.

II. EXPERIMENTAL AND THEORETICAL METHODS

A. Experimental data

The temperature-dependent magnetic susceptibility of a powdered sample of Al_2Er_2 has been reported in Ref. [57]. For convenience, the χT vs T curve is replotted in Fig. 2(b). The molecule exhibits a butterfly structure which contains two magnetic Er^{III} ions and two diamagnetic Al^{III} ions as its metal centers [Fig. 2(a)]. The lack of exchange coupling and negligible dipolar coupling between the Er^{III} ions means that they

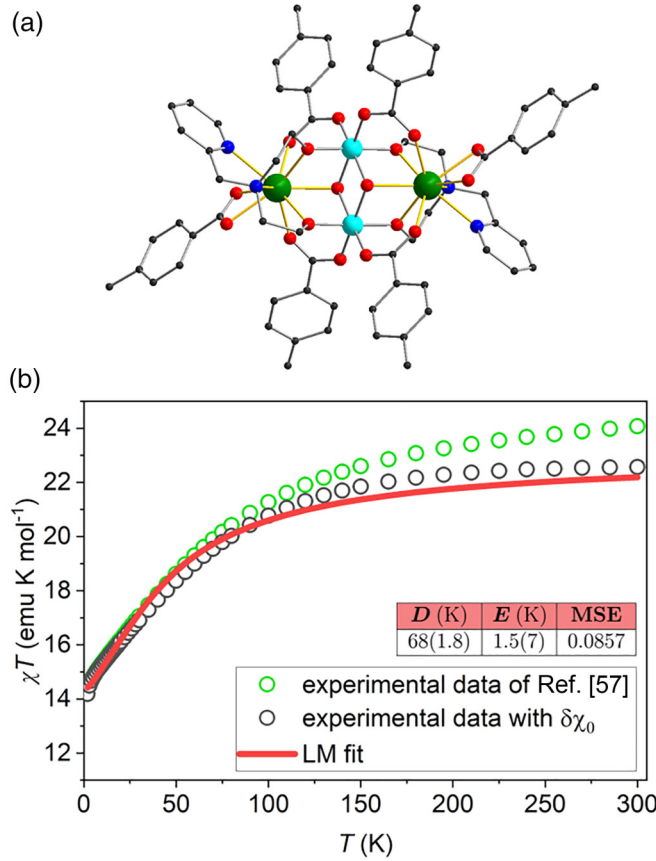


FIG. 2. (a) Molecular structure of the butterfly molecule Al_2Er_2 . The diamagnetic Al^{III} ions (cyan) form the body, whereas the magnetic Er^{III} ions (green) form the wings of the butterfly structure (O: red, N: blue, C: black, H atoms are omitted). (b) LM fit to the experimental susceptibility curve of Al_2Er_2 . For reference, the experimental data of Ref. [57] are shown (green open circles). The experimental curve used in this work involves an empirically determined adjustment of the diamagnetic contribution, as described in the text (gray open circles) ($\delta\chi_0 = -0.005 \text{ emu mol}^{-1}$, $m_c = 1$). The red solid line represents the simulated curve corresponding to the LM fit parameters (the parameter values are given in the text).

can be modeled as independent single ions. This property as well as the diamagnetic behavior of the Al^{III} makes Al_2Er_2 an excellent test-bed to allow an isolated study of $4f$ ligand field parameters. The total angular-momentum quantum number of Er^{III} is $J = 15/2$, the Landé g factor is $g_J = 1.2$ [57].

The experimental data of Ref. [57] have been properly treated, i.e., corrected for diamagnetic contributions from the sample holder and other contributions from the experimental setup and scaled according to the measured sample mass (Pascal's constants [58] were not applied). However, generally, except in especially carefully conducted experiments, uncertainties remain; for instance, the diamagnetic correction or the determination of the molar mass may not be perfectly precise. Given the featureless nature of the magnetic curves, these uncertainties can significantly affect the values of parameters obtained from, e.g., least-squares fits to the data. It is therefore generally advisable to include these experimental uncertainties in the modeling and investigate their effect on the determined parameter values. In this work the two most

relevant experimental uncertainties are considered, which are a correction $\delta\chi_0$ accounting for inaccuracies in the estimation of the diamagnetism, and a molar mass correction factor m_c , which accounts for inaccuracies in the measured sample mass (in the ideal case their values are $\delta\chi_0 = 0$ and $m_c = 1$). On an empirical basis, the experimental data of Ref. [57] was corrected with $\delta\chi_0 = -0.005 \text{ emu mol}^{-1}$ [see Fig. 2(b)] and these data are used henceforth in all analyses.

B. Magnetic models

Increasingly more complex contributions to the ligand field Hamiltonian $\hat{H}_{\text{LF}}$, Eq. (1), are considered in a stepwise manner, starting with the second-order ($k = 2$) terms. Up to second order, the ligand field Hamiltonian $\hat{H}_{\text{LF}}$ can be written as

$$\sum_{q=-2}^2 B_{2q} \hat{O}_{2q} = \hat{\mathbf{J}}^T \begin{bmatrix} B_{2,2} - B_{2,0} & B_{2,-2} & \frac{1}{2}B_{2,1} \\ B_{2,-2} & -B_{2,2} - B_{2,0} & \frac{1}{2}B_{2,-1} \\ \frac{1}{2}B_{2,1} & \frac{1}{2}B_{2,-1} & 2B_{2,0} \end{bmatrix} \hat{\mathbf{J}}, \quad (2)$$

where $\hat{\mathbf{J}}$ is the total angular-momentum vector operator. This equation can also be written as $\hat{\mathbf{J}}^T \mathbf{D} \hat{\mathbf{J}}$, where the tensor $\mathbf{D}$ is implicitly given by Eq. (2). The tensor $\mathbf{D}$ is real, symmetric, and traceless, which implies that it is always possible to rotate the coordinate frame so that $\mathbf{D}$ is diagonal. In such a frame then,

$$\hat{\mathbf{J}}^T \mathbf{D} \hat{\mathbf{J}} = \hat{\mathbf{J}}^T \begin{bmatrix} -\frac{1}{3}D + E & 0 & 0 \\ 0 & -\frac{1}{3}D - E & 0 \\ 0 & 0 & \frac{2}{3}D \end{bmatrix} \hat{\mathbf{J}}. \quad (3)$$

As this corresponds to aligning one of the three coordinate axes along the principal magnetic anisotropy axis of the molecule, there exist three physically equivalent parameter sets D, E corresponding to each choice of coordinate axis. Choosing the z axis imposes the condition $|E| \leq 1/3|D|$. The other two parameter sets are then given by

$$\begin{aligned} D' &= -\frac{1}{2}D - \frac{3}{2}E, & E' &= +\frac{1}{2}D - \frac{1}{2}E, \\ D'' &= -\frac{1}{2}D + \frac{3}{2}E, & E'' &= -\frac{1}{2}D - \frac{1}{2}E. \end{aligned} \quad (4)$$

After specifying the z axis, there remains a choice in the directions of the x and y axis, an interchange of which leads to a minus sign in the E term. The three parameter sets resulting from the choice of z axis combined with the indeterminacy of the sign of E lead to six descriptions in the space of second-order ligand field parameters, which are physically equivalent in that their associated powder magnetic susceptibility curves are identical (when only D and E are included in the model). The parameter set with $|E| \leq 1/3|D|$ and $E \geq 0$ will be denoted as the “standardized parameter set.”

Initially, this work will consider a second-order magnetic model in a nonstandardized range of D and E parameters, which provides a trivial but excellent test case of determining multiple physically equivalent solutions for a given susceptibility curve. This model will be referred to as the two-parameter model. The model is then extended to include one additional fourth-order ligand field parameter in an attempt to gradually increase over-parametrization. Specifically,

the terms $B_{4,4}$ and $B_{4,-4}$ will be discussed. These models will be referred to as three-parameter models. Finally, a five-parameter model will be considered which includes $B_{4,0}$, $B_{4,1}$, and $B_{4,-4}$ in addition to the D and E parameters. In the rest of this paper, the second-order ligand field parameters will be represented in the D and E notation, while the higher-order parameters will be given in the B_{kq} notation.

C. Variational autoencoder and invertible neural network

A comprehensive explanation of VAEs can be found in Ref. [59]. The VAE is employed in this work to encode the susceptibility data into a low-dimensional latent space representation. This machine-learning model achieves this by maximizing the evidence lower bound [59] which corresponds to the minimization of the loss $\mathcal{L}_{\text{VAE}}$,

$$\mathcal{L}_{\text{VAE}} = (\chi T - \widehat{\chi T})^2 - \frac{1}{2}\beta(1 + \sigma_y - \mu_y^2 - e^{\sigma_y}), \quad (5)$$

where the first term is the reconstruction loss as implemented by a mean squared error and the second term represents the Kullback-Leibler divergence [59], where μ_y and σ_y are the mean and variance of the encoded latent space vector $\mathbf{y}$. The prefactor β is used as a weight to balance the disentangling in the latent dimensions against the reconstruction accuracy [60].

The objective of the INN is to estimate the posterior distribution $p(\mathbf{x}|\mathbf{y})$, where $\mathbf{x}$ represents (possibly multiple) sets of ligand field parameters and $\mathbf{y}$ is the latent space representation of the susceptibility data as encoded by the VAE. The INN addresses the inverse problem by introducing additional random variables sampled from a Gaussian distribution. This approach creates a bijective mapping which is represented by the deterministic function $\mathbf{x} = g(\mathbf{y}, \mathbf{z}; \theta)$, with $p(\mathbf{z}) = N(\mathbf{z}; 0)$, where $\mathbf{z}$ is a set of Gaussian-sampled variables and θ represents the parameters of the INN. The invertible architecture allows simultaneous learning of the inverse process $g(\mathbf{y}, \mathbf{z}; \theta)$ as well as an approximation $f_y(\mathbf{x}; \theta)$ of the forward process according to [56]

$$[\mathbf{y}, \mathbf{z}] = f(\mathbf{x}; \theta) = [f_y(\mathbf{x}; \theta), f_z(\mathbf{x}; \theta)] = g^{-1}(\mathbf{x}; \theta). \quad (6)$$

The models f and g share the same parameters θ as implemented in a single INN. The posterior distribution which the network models is then given by

$$p(\mathbf{x} = g(\mathbf{y}, \mathbf{z}; \theta) | \mathbf{y}) = p(\mathbf{z}) | J_x |^{-1},$$

$$J_x = \det \left(\frac{\delta g(\mathbf{y}, \mathbf{z}; \theta)}{\delta [\mathbf{y}, \mathbf{z}]} \right), \quad (7)$$

where J_x is the Jacobian determinant which can be calculated using real-valued non-volume preserving (RNVP) coupling blocks [61]. The overall loss to be minimized during the INN training is given by

$$\mathcal{L}_{\text{INN}} = \lambda_x \mathcal{L}_x + \lambda_y \mathcal{L}_y + \lambda_z \mathcal{L}_z, \quad (8)$$

where $\mathcal{L}_y = (\chi T - \widehat{\chi T})^2$ is the mean squared error (MSE) reconstruction loss and $\mathcal{L}_x$ and $\mathcal{L}_z$ are maximum mean discrepancy [62] losses used to fit the probability distributions of the prior and the Gaussian, respectively. The full details

of INNs, along with comparisons of performance to other machine-learning models, can be found in Ref. [56]. The MSE corresponds to the χ^2 value in the context of least-squares fitting.

D. Training

For the second-order ligand field parameters D and E a nonstandardized range of values was considered in all experiments (except when stated otherwise). For the two-parameter model the dataset contained 40 402 simulated susceptibility curves in the parameter range $D \in [100, -100]$ K, $E \in [-50, 50]$ K. The three-parameter datasets involving, in addition, one fourth-order parameter contained 564 000 susceptibility curves in the range $D \in [-100, 100]$ K and $E \in [-50, 50]$ K. As preprocessed input to the VAE-INN, all susceptibility data was normalized to lie within $[0, 1]$ by the transformation

$$\chi T = \frac{\widetilde{\chi T} - \widetilde{\chi T}_{\min}}{\widetilde{\chi T}_{\max}}, \quad (9)$$

where $\widetilde{\chi T}$ is the original data from the dataset. An 80/20 split was made between the training and testing datasets for the purpose of validation.

The encoder component of the VAE consisted of an input layer of 300 nodes and 3 hidden layers of 80, 20, and 8 nodes. Tangent hyperbolic activation was used. The β factor in Eq. (5) was set to 5×10^{-3} in order to prioritize reconstruction accuracy, which was critical to accurate parameter predictions. A learning rate of 1×10^{-3} was used along with a batch size of 300. At this configuration, 800 epochs were found to be sufficient for a well-generalized performance on the test data. ADAM was chosen as gradient optimizer. Optimal dimensionality of the latent space was determined by reducing it until the loss on the test-set increased dramatically. This has to be assessed on a case-by-case basis for each dataset because each new ligand field parameter may introduce significantly distinct variations which may require additional dimensions of the latent space. The MSE losses for varying latent space dimensionality are shown in Fig. S1 of the Supplemental Material [63]. A latent space of two dimensions was found to successfully encode the two-parameter dataset and had to be increased to three dimensions in order to encode the data simulated from higher-order ligand field parameters.

The INN training was found to be much more successful and efficient if trained not upon the entirety of the VAE latent space but a local region surrounding the encoding of experimental data. The radius of this region had to be set as a hyper-parameter dependent on the latent space dimensions. The INN consisted of four RNVP coupling blocks with hidden layer sizes of 64 nodes. Tangent hyperbolic activation was found to be robust at maintaining the expected range of parameter predictions. The weighting factors for the losses in Eq. (8) were set as $\lambda_y = 1$, $\lambda_z = 40$, and $\lambda_x = 250$. With a batch size of 64, the forward training was performed at a learning rate of 5×10^{-4} for 2000 epochs while the backward training was performed at a reduced learning rate of 5×10^{-5} for up to 6000 epochs depending on the complexity of the magnetic model. The ADAM optimizer was used.

E. Computational details

The magnetic susceptibility curves were simulated using in-house code. The machine-learning code was written in Python and used PYTORCH for the machine-learning models, while the relevant preprocessing of the susceptibility dataset utilized the Python libraries PANDAS and NUMPY. The training was conducted on an NVIDIA RTX 4080 GPU with 16 GB of memory.

III. RESULTS AND DISCUSSION

The aim of the following experiments is to determine the ability of the VAE-INN to predict ligand field parameters in situations where the characterization of the magnetic susceptibility data is nonunique and over-parametrized. The sixfold invariance of the D , E parameters in the two-parameter model provides a test case to tackle a generic one-to-many problem in the space of the second-order ($k = 2$) ligand field parameters. The analysis is then extended to models which include a few fourth-order parameters. Confidence in the VAE-INN predictions is in addition established by testing the VAE-INN on simulated susceptibility curves, the ligand field parameters for which are thus known and can be compared with the VAE-INN results.

To compare the accuracy of the VAE-INN model with a benchmark, a LM fit was performed on the experimentally measured susceptibility curve of Al_2Er_2 [57]. The diamagnetic and molar mass corrections, $\delta\chi_0$ and m_c , can significantly alter the LM best fit parameters and thus were also considered in addition to the D , E parameters, but this further contributes to over-parametrization. The resulting fit to the experimental susceptibility curve and the corresponding best-fit parameters are shown in Fig. 2(b) ($m_c = 1$ was presumed).

The D and E parameters were determined to be 68(1.8) K and 1.5(7) K, respectively. It is important to emphasize that iterative numerical methods like LM require a good initialization point on the parameter landscape in order to converge into a nearby minimum, which may not be the global minimum [30]. To achieve convergence in over-parametrized situations, as encountered here, it can be practical to use a trial and error approach and to first obtain a converging best-fit value for a single parameter before fitting the next parameter, and so on. This stepwise, human-guided approach had to be implemented in the present case in order to get the LM algorithm to converge to a result. The best-fit parameters were found to be highly dependent upon initialization values as well as the chosen sequencing of parameters. In the absence of a ground truth for the parametrization of the Al_2Er_2 data, the best-fit parameters in Fig. 2 were taken as benchmarks for comparison with predictions of the VAE-INN model.

A. Two-parameter dataset

As detailed before, the nonstandardized D and E parameter range can exhibit six equally valid parameter sets which describe identical susceptibility curves. Typically, these parameters would be considered in their standardized ranges so that the trivial symmetry resulting from the choice of coordinate axes is ignored. However, the nonstandardized range presents us with a multiplicity of solutions which can be used

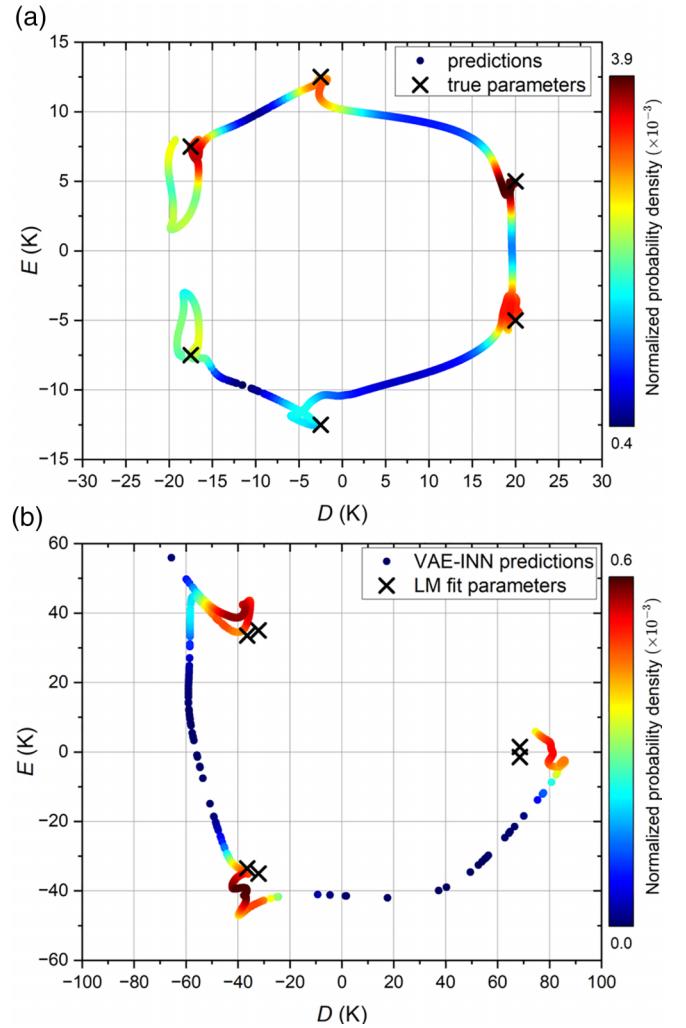


FIG. 3. VAE-INN predictions in the D and E parameter space. The VAE-INN was used to sample 10000 parameter predictions which are plotted as solid circles and colored according to their normalized probability density (red: highest probability, dark blue: lowest probability). (a) Predictions for a simulated susceptibility curve with $D = 20$ K and $E = 5$ K. The black crosses mark the six physically equivalent parameter sets corresponding to these D , E values. (b) Predictions for the experimental Al_2Er_2 susceptibility curve. The black crosses mark the six physically equivalent parameter sets which correspond to the parameters obtained from the LM fitting of this curve [its D , E values are given in the text, and also in Fig. 4(a)].

as an artificial test case for the more general situation where solutions are not physically equivalent. A simple inverse problem can therefore be created by training the VAE-INN model on the range of parameter values of $D \in [100, -100]$ K, $E \in [-50, 50]$ K. As an initial test case, the VAE-INN model is tasked to predict the parameters for a simulated susceptibility curve for which the ground truth parameter values are thus known. In the following case, the susceptibility curve was simulated from a Hamiltonian with $D = 20$ K and $E = 5$ K. Figure 3(a) shows 10000 predictions of the INN (after the VAE-INN was trained) as well as the six known solutions, which correspond to the D , E values used in the simulated susceptibility curve. The parameter predictions are colored

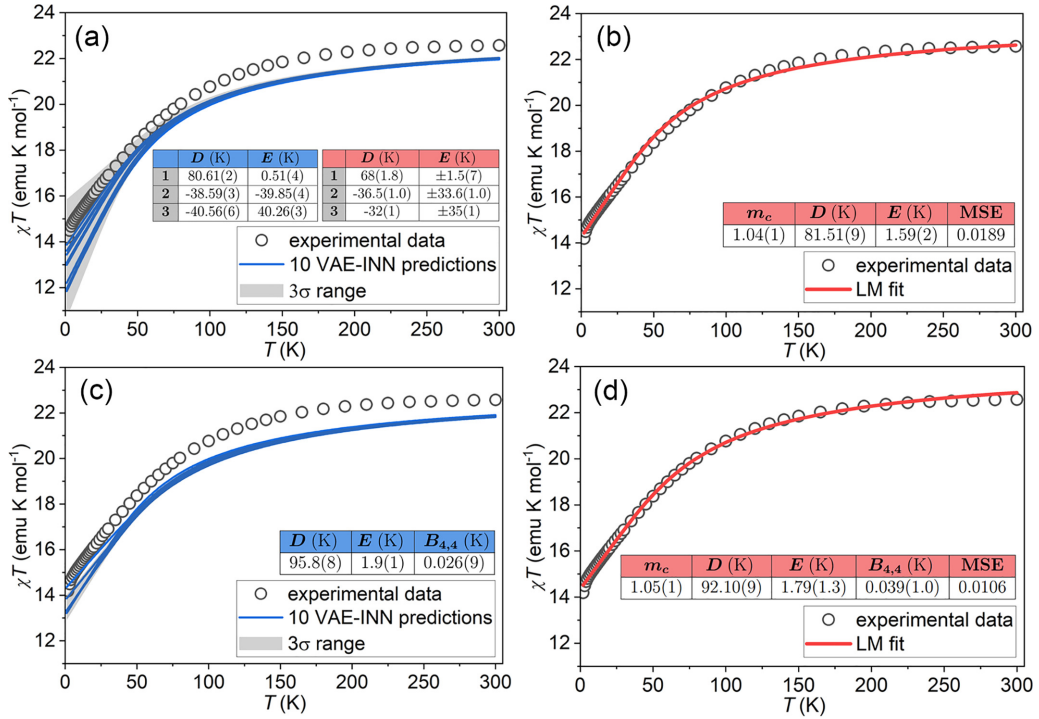


FIG. 4. VAE-INN predictions (blue solid lines) and LM fits (red solid lines) for ligand field parameters. In all panels, open circles represent the experimental susceptibility curve of Al_2Er_2 . (a) Two-parameter model: Simulated susceptibility curves corresponding to 10 VAE-INN predictions for D and E parameters shown as blue solid lines. The variability among all predicted curves is shaded gray in a range of three standard deviations. The mean of each predicted parameter set and the standard error are listed in the left (blue) table for VAE-INN, whereas the right (red) table lists the six physically equivalent parameter sets which correspond to the parameter set obtained from LM fitting. (b) Two-parameter model: LM fit to the parameters m_c , D , and E shown as red solid line. The best-fit parameters are listed in the table. (c) Three-parameter model: Simulated susceptibility curves corresponding to 10 VAE-INN predictions for D , E , and $B_{4,4}$ parameters shown as blue solid lines. The variability among all predicted curves is shaded gray in a range of three standard deviations. The mean of each predicted parameter and the standard error are listed in the table. (d) Three-parameter model: LM fit to the parameters m_c , D , E , and $B_{4,4}$ shown as red solid line. The best-fit parameters are listed in the table.

according to their normalized probability density in the parameter space.

Six regions of high prediction density can clearly be seen, which closely match the six expected D , E solutions. Contrary to expectation, the normalized probability density is not evenly distributed among the six solutions. Upon training the model multiple times, the dominant solutions were found to vary randomly, indicating a symmetry breaking of the model's optimizer resulting from random weight initialization [64]. The symmetry can be recovered by averaging over the outputs of multiple trained models. This means that the normalized probability density in one of the trained models does not correlate with the MSE between the predicted and simulated susceptibility data. Instead, the model only predicts parameter solutions in distinct regions with no indication of their relative preference.

Next, the VAE-INN has been used to predict D and E parameter values for the experimental Al_2Er_2 susceptibility curve. Figure 3(b) shows the VAE-INN predictions compared with the six D , E parameter sets corresponding to the LM fit solution. The parameter space contains three concentrated regions containing a high probability density of predictions. In this case, only three instead of six solutions appear to be present since each two of them are not well separated due to the low value of E in comparison to D . In general,

it was found that the VAE-INN predictions have a certain resolution which may depend on the landscape of solutions in the parameter space. While pockets of solutions further apart are identified, as in Fig. 3(a), relatively close-lying solutions are not resolved. Also the span in a parameter value of the identified solutions sets a kind of resolution for the detectable differences in this parameter. That is, existing solutions with small differences might not be identified in the presence of solutions with large differences. The model's uncertainty on the predicted solutions, in contrast, can be quantified by the standard error of the mean (or standard error in brief), which is given by $\sigma/\sqrt{n}$, where σ is the standard deviation for the considered solution and n is the number of samples. The largest standard error among the three high-density regions was found to be 0.06 and 0.04 K for D and E , respectively. These regions lie very close to the LM fit solutions, and hence the six characterizing parameter sets have been successfully captured by the VAE-INN model within its resolution. The mean value for each of the predicted parameter sets are, however, somewhat shifted towards larger magnitude compared with the LM best-fit parameters. The result of this shift can also be seen in the simulated susceptibility curves for the predicted parameters shown in Fig. 4(a), which are generally lower in value than the experimental curve. This suggests that the observed shift in the predictions may be caused by a

molar mass correction m_c , which scales the susceptibility with a multiplicative factor. To test this hypothesis, the VAE-INN parameter predictions were used as an initial guess for a LM fit performed on D , E , and m_c . The resulting fit is shown in Fig. 4(b). In this case, the LM fit found that $D = 81.51(9)$ K, $E = 1.59(2)$ K and $m_c = 1.04(1)$. Notably, the best-fit parameters for D and E match now those predicted by the VAE-INN in Fig. 4(a). Therefore, as a conclusion of the findings so far, the VAE-INN is able to determine the D , E parameters with good accuracy without explicitly minimizing the MSE between predicted susceptibility curves and experimental data, therefore ignoring any shifts caused by a molar mass correction. The small difference in the parameter predictions as compared with the LM benchmark can be associated with a molar mass correction not taken into account in the LM benchmark fit.

The VAE-INN model offers significant advantages over LM fitting. First, the VAE-INN not only predicted six sets of D and E parameters which fit the experimental data, but also significantly improved upon the initial LM benchmark results in Fig. 2(b). That is, it could determine more realistic parameter values than LM without having been biased towards molar mass corrections. This improvement can be associated with the preprocessing of the training dataset shown in Eq. (9), which scales all curves to be in the range $[0,1]$. Such a preprocessing renders the model insensitive to any magnitude scaling of the data, resulting in the parametrization of curvature only. The proposed VAE-INN is therefore relatively robust against molar mass corrections, which was required as a necessary fit parameter for LM to produce an equivalent result. Furthermore, due to the prior training of the VAE-INN model on a specified range of ligand field parameters, it is able to seek an MSE minimum with a convergence that is independent of any initialization within the entire range. This also allowed the VAE-INN to predict parameters simultaneously, in contrast with LM which required human-guided convergence through sequential parameter fitting.

B. Three-parameter datasets

The fit to the experimental data may potentially be improved by introducing higher-order ($k \geq 4$) ligand field parameters. The choice and range of parameters to consider, however, is *a priori* unknown. The parameters may also be nonlinearly correlated, which leads to a highly nontrivial overparametrization. Again, as an initial test, the VAE-INN was first tasked with predicting D , E and an additional significant fourth-order parameter for simulated susceptibility curves. A fourth-order parameter is considered significant when it alters the energy spectrum, and hence the susceptibility, in a manner not achievable by only varying the D and E parameters. For the $B_{4,4}$ term, this is for instance realized for $D = 14$ K, $E = 39$ K, and $B_{4,4} = 0.37$ K. A similar situation is obtained for the $B_{4,-4}$ term when $D = 78$ K, $E = 1$ K, and $B_{4,-4} = 0.19$ K. Both parameter sets create two low-lying Kramers doublets in the energy spectrum which are close in energy, and which lead to a significant drop of the susceptibility at low temperatures. The simulated susceptibility curves for the two parameter sets are shown in Fig. S2 in the Supplemental Material [63]. In the following, the two examples of higher-order parameter

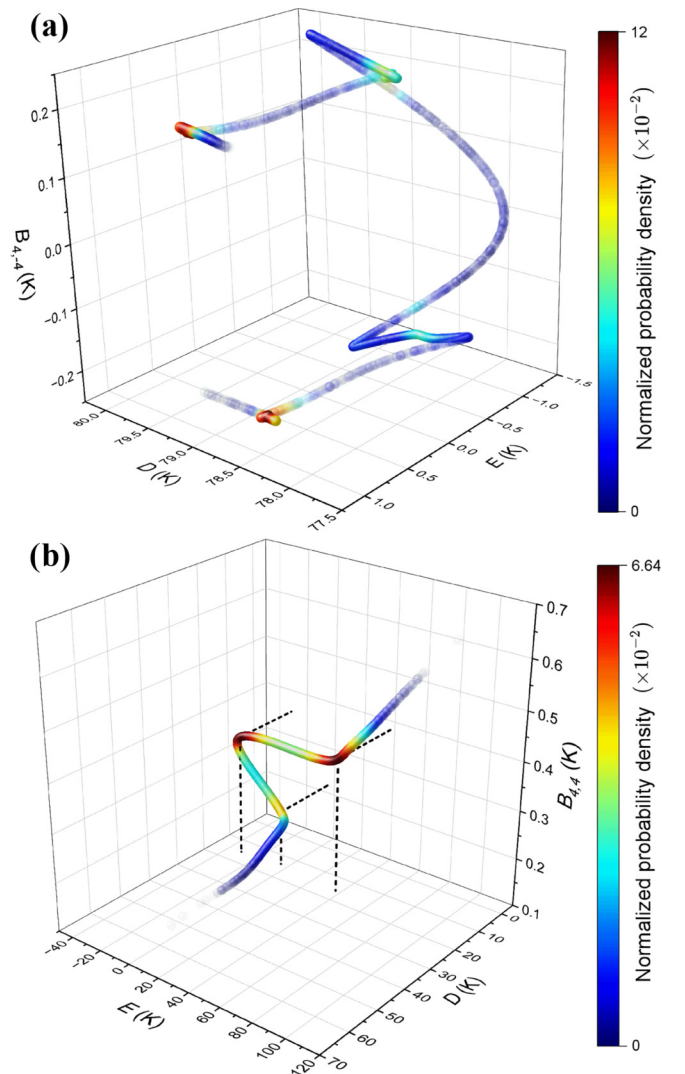


FIG. 5. VAE-INN predictions in the three-parameter spaces discussed in the text for simulated susceptibility curves. The VAE-INN was used to sample 10 000 parameter predictions, which are plotted as scatter points and colored according to their normalized probability density. (a) VAE-INN predictions for the simulated susceptibility curve with $D = 78$ K, $E = 1$ K, and $B_{4,-4} = 0.19$ K. (b) VAE-INN predictions for the simulated susceptibility curve with $D = 14$ K, $E = 39$ K, and $B_{4,4} = 0.37$ K. The dotted lines represent the projections of the high-density regions onto the D , E and E , $B_{4,4}$ parameter planes, for better visualization of the associated parameter values.

contributions are treated separately. Note that the arguments which lead to six equivalent parameter sets in the D , E parameter space do not apply here as such anymore. The indeterminacy in the sign of E remains but otherwise any further multiplicity, if any, needs a case to case analysis.

Two datasets of 564 000 susceptibility curves were created with the range of D , E parameter values as before and parameters $B_{4,4} \in [-0.5, 0.5]$ K or $B_{4,-4} \in [-0.5, 0.5]$ K, respectively. Figure 5 shows the VAE-INN predictions in these three-dimensional parameter spaces for the respective simulated susceptibility curve. The predictions shown in Fig. 5(a) are for the D , E , and $B_{4,-4}$ model, and four areas of high density are observed for parameter values of $D \approx 79$ K,

$E \approx \pm 0.7$ K, and $B_{4,-4} \approx \pm 0.2$ K (the precise values are given for each area in Table S1 in the Supplemental Material [63]). The predictions for the D , E , and $B_{4,4}$ model are presented in Fig. 5(b), where two areas of high-density are seen for parameter values of $D \approx 16$ K, $E \approx \pm 35$ K, and $B_{4,4} \approx 0.38$ K (the precise values are given for each area in Table S2 in the Supplemental Material [63]). It can be observed that the VAE-INN successfully predicts the known parameter solutions in both cases, including some sign indeterminacy. In particular, the VAE-INN model identifies a sign indeterminacy in $B_{4,-4}$ which is not present in $B_{4,4}$. This has been confirmed by independent explicit numerical simulations. In the D , E , and $B_{4,4}$ model, an additional third region of lower but significant probability is seen at approximately $D \approx 13$ K, $E \approx -7$ K, and $B_{4,4} \approx 0.22$ K. Although relatively inaccurate at the lowest temperatures, this parameter set also describes the simulated input curve well (see Fig. S3 in the Supplemental Material [63]) and adequately captures the overall susceptibility behavior. The expected sign indeterminacy of E appears to not be observed here, which is, however, again associated with the resolution limit of the VAE-INN in this case, i.e., to the fact that the E value is considerably smaller than in the other two solutions found for this model. Regardless, the presence of the third solution indicates a capability of the VAE-INN to discover new parameter sets which are not simply related mathematically as in the case of the D , E model. It can be concluded that the VAE-INN model retains its ability to predict regions of nonunique solutions in the space of three parameters.

Next the experimental Al_2Er_2 susceptibility curve was supplied to the VAE-INN, trained for each of the two three-parameter datasets. The LM fit achieved in Fig. 4(b) for the two-parameter model does not necessitate the presence of fourth-order ligand field contributions. Hence, any fit to a $B_{4,q}$ term is likely to be overfitting. Nevertheless, the performance of the VAE-INN on the Al_2Er_2 data in the presence of such an overfitting parameter can be investigated. As a semi-arbitrary choice the $B_{4,4}$ parameter was considered. To focus on the effect of the fourth-order parameter, a standardized range of D and E parameters was considered in this case, thereby eliminating the sixfold multiplicity in these parameters and a corresponding multiplicity in the solutions for small values of $B_{4,4}$. A plot of the VAE-INN predictions are shown in Fig. S4 in the Supplemental Material [63]. Notably, two regions of solutions are observed with slightly different probabilities. The one with the slightly lower probability essentially corresponds to $B_{4,4} = 0$ and thus essentially recovers the result of the two-parameter model (see Fig. S4 in the Supplemental Material [63]). For the region with the higher probability, Fig. 4(c) shows the simulated susceptibility curves corresponding to 10 predictions of D , E , and $B_{4,4}$ parameters. The standard deviation among the susceptibility curves is greatly reduced in the case of predicting a single parameter set as solution. The predicted D and E parameters were $D = 95.8(8)$ K and $E = 1.9(1)$ K, while $B_{4,4}$ was predicted to be $0.026(9)$ K. The D and E values are slightly larger than those obtained from the two-parameter model. The predicted susceptibility curves are consistently lower in value than the experimental curve, similar to the earlier observation for the two-parameter case, where it could be attributed to a molar mass correction.

To obtain an LM fit of parameters m_c , D , E , and $B_{4,4}$ to the experimental data, a human-guided fitting procedure had again to be used. The LM fit converges to a similar minimum as found by the VAE-INN, but its successful convergence is subject to the condition that the initial values of the parameters in each step are not far off. The result of this fit and the determined LM best-fit parameters are shown in Fig. 4(d). The LM fit found similar ligand field parameters with an additional molar mass correction $m_c = 1.05(1)$. By comparing the MSE in Figs. 4(b) and 4(d), it can be seen that the introduction of $B_{4,4}$ slightly improves the fit to experimental data. A similar improvement, however, was also obtained with $B_{4,-4}$ instead of $B_{4,4}$, which indicates that this improvement is not significant or unique. Regardless, it can be concluded that the VAE-INN model accurately predicts up to three ligand field parameters that reduce the MSE between simulated and experimental susceptibility data without having to incorporate a molar mass correction. Furthermore, it offers a much more robust initialization compared with LM fitting, of course only as long as the MSE minimum lies in a parameter space region which is spanned by the training dataset.

C. Five-parameter dataset

To investigate the VAE-INN's performance on a higher-dimensional parameter space, the dataset was extended to simultaneously include the $B_{4,0}$, $B_{4,1}$, and $B_{4,-4}$ parameters, each in the range $[-0.4, 0.4]$ K, in addition to the D and E parameters. A standardized range of D and E parameter values was again considered. The dataset of 2.3 million curves generated from this five-dimensional parameter space was successfully encoded by the VAE into a three-dimensional latent space. This indicates that the distinct type of variations in the susceptibility curves resulting from the five ligand field parameters can in fact be reduced to the effect of three hidden parameters; a redundancy which is to be expected in an overparametrized model. No significant information is lost in this encoding since allowing more than three hidden parameters does not improve the VAE performance (see Fig. S1 in the Supplemental Material [63]). The INN, in contrast, has been observed to be significantly more difficult to train on this higher-dimensional space of parameters, requiring an extensive fine-tuning of the model's hyper-parameters. Figure 6(a) shows the obtained VAE-INN parameter predictions as projected onto the plane spanned by the D and E parameters. Projections for the $B_{4,0}$, $B_{4,1}$, and $B_{4,-4}$ parameter predictions are provided in Fig. S5 in the Supplemental Material [63].

Some points are to be noted. First, the VAE-INN found only a single solution in this space of five parameters with the parameter values given in Fig. 6(b). The value of D , predicted to $D \approx 50$ K, is significantly lower compared with all previously discussed models. This can be associated with the inclusion of a $B_{4,0}$ parameter, which now acts in combination with D . Regardless, the VAE-INN should have been expected to also identify the previously known good solutions with $B_{4,0} = 0$. Furthermore, the precision of the predictions is greatly reduced. This can also be seen in the corresponding simulated susceptibility curves. Ten representative VAE-INN parameter predictions are shown in Fig. 6(b), along

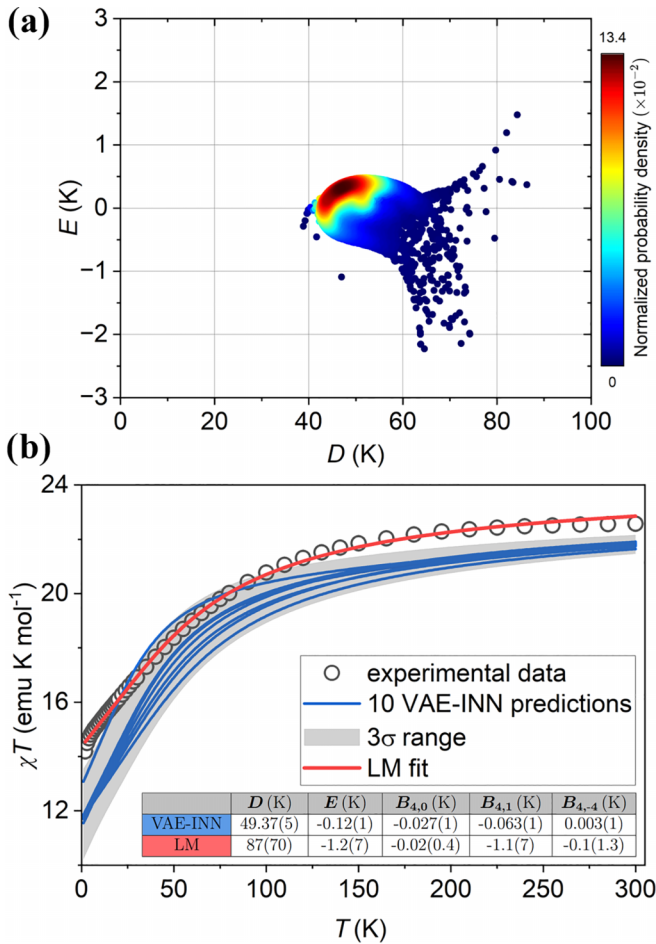


FIG. 6. VAE-INN predictions for the experimental susceptibility curve of Al_2Er_2 using the five-parameter model involving the D , E , $B_{4,0}$, $B_{4,1}$, and $B_{4,-4}$ parameters. (a) VAE-INN predictions projected onto the space of D and E parameters. The VAE-INN was used to sample 10 000 parameter predictions which are plotted as scatter points and colored according to their normalized probability density. One area of high density is visible. (b) Susceptibility curves corresponding to 10 sets of predictions for D , E , $B_{4,0}$, $B_{4,1}$, and $B_{4,-4}$ (blue solid lines). The variability among all predicted curves is shaded gray in a range of three standard deviations. Open circles represent the experimental susceptibility curve. A LM fit to the experimental data for the five parameters is also shown (red solid line). The predicted parameters by both VAE-INN and LM fit are listed in the table.

with the experimental Al_2Er_2 curve. The loss of precision in parameter predictions is reflected in the greater variability of the simulated susceptibility curves. As before, the parameter values predicted by VAE-INN can be used as initial values for an LM fit, the result of which is also shown on the plot. The error ranges on all the best-fit parameter values are on the same order of magnitude as the values themselves, which indicates a high degree of overfitting and a general limit of using LM fits with a large number of parameters. The VAE-INN model, while still predicting parameter values that reproduce the curvature of the experimental curve reasonably well, also begins to struggle in this five-parameter space. It loses its ability of predicting all ligand field parameter sets in the given

five-dimensional space which could describe the data equally well. Moreover, the precision of the predictions is reduced.

Taking higher-order ligand field parameters into account rapidly increases the size of the training dataset, and the parameter space cannot be densely covered in all areas. It indeed was found in general that the INN struggled to learn from this high-dimensional parameter space as the region of interest surrounding the VAE encoding of the experimental susceptibility curve was sparsely populated. This issue is commonly known as the curse of dimensionality [65]. While more data may be easily simulated for a region of interest, the local regions in parameter space relevant to the input data are unknown *a priori* and the vast majority of data elsewhere does not benefit the learning of the VAE-INN. This causes an exponentially increasing inefficiency of learning that is, for the INN, significant already at five ligand field parameters in the present case, and motivates the need for more intelligent navigation and sampling in the parameter space.

IV. CONCLUSIONS

In this work, a variational autoencoder and an invertible neural network were combined, and the resulting VAE-INN architecture was demonstrated to be able to predict ligand field parameters for measured powder magnetic susceptibility curves of SMMs with an accuracy comparable to that of least-squares fitting. Importantly, this data-driven approach has been found to offer notable advantages such as a robust initialization process, provision of nonunique solution predictions, and the capability to generalize over molar mass inaccuracies inevitable in experiments. This capacity for generalization enables the proposed VAE-INN model to successfully predict D and E values that more accurately characterized the experimental data compared with least-squares best-fit results. The VAE-INN model has been found to maintain its predictive capacity for the prediction of up to three ligand field parameters as well as being insensitive to the molar mass correction parameter. As least-squares fitting is sensitive to this experimental uncertainty, it has explicitly to be included in the fitting in order to achieve a similar description in terms of the ligand field parameters. The VAE-INN model has also been found to be comparatively robust within the parameter space it was trained on, whereas the convergence of the least-squares fitting was highly dependent upon the initialization values within the same parameter space. This advantage becomes increasingly substantial as more parameters are considered.

As the VAE-INN approach is strictly data driven, however, it is vulnerable to approaching a limit due to the sparsity of the simulated training dataset, which in the present use case has been found to be reached for a five-dimensional space of ligand field parameters. In future work, the onset of such a limitation might be significantly delayed through simulation embedded algorithms that focus on further populating local regions of interest in the latent space or Markov chain Monte Carlo methods that can be used to approximate a more informative prior distribution in the parameter space [66,67]. In addition, integrating known information into the machine-learning model or physics-informed decisions in the preprocessing, architectural design and loss functions can

potentially lead to far more efficient usage of data and thus accessing of significantly larger parameter spaces [68,69].

Another obvious path for improvement is to include additional experimental data such as field dependent powder magnetization curves at low temperatures, which are also routinely measured. However, the numerical simulation of such curves is massively more time consuming (in comparison with the simulation of the magnetic susceptibility) since a powder average has to be performed. This limits the dataset sizes quite substantially, and more efficient strategies for populating the relevant regions are thus mandatory. It needs to be emphasized however, that the long simulation times for these curves also severely limits the least-squares techniques. In the long term, machine-learning methods may have an edge over least-squares techniques, since for a given parameter space the dataset needs to be created only once, whereas in the least-squares methods the magnetization curves need to be simulated on the fly.

The methodology employed in the VAE-INN framework as pursued in this work can complement traditional numerical optimization methods within the realm of SMMs. Its

success in predicting nonunique parameter solutions, and the additional benefits mentioned in the above, can open up a data-driven approach to exploring not only the significance of higher-order ligand field parameters but also the correlations between them. Such information on the nature and correlation of higher-order parameters can help provide insights into the elusive ligand field parametrization of lanthanides.

ACKNOWLEDGMENT

The authors thank the group of A. K. Powell for providing the experimental data of Al_2Er_2 in electronic format.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available upon publication because it is not technically feasible and/or the cost of preparing, depositing, and hosting the data would be prohibitive within the terms of this research project. The data are available from the authors upon reasonable request.

- [1] G. Carleo and M. Troyer, Solving the quantum many-body problem with artificial neural networks, *Science* **355**, 602 (2017).
- [2] X. Gao and L.-M. Duan, Efficient representation of quantum many-body states with deep neural networks, *Nat. Commun.* **8**, 662 (2017).
- [3] H.-Y. Huang, R. Kueng, G. Torlai, V. V. Albert, and J. Preskill, Provably efficient machine learning for quantum many-body problems, *Science* **377**, eabk3333 (2022).
- [4] 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, *Commun. Mater.* **3**, 84 (2022).
- [5] A. M. Samarakoon, P. Laurell, C. Balz, A. Banerjee, P. Lampen-Kelley, D. Mandrus, S. E. Nagler, S. Okamoto, and D. A. Tennant, Extraction of interaction parameters for RuCl_3 from neutron data using machine learning, *Phys. Rev. Res.* **4**, L022061 (2022).
- [6] K. T. Schütt, H. Glawe, F. Brockherde, A. Sanna, K. R. Müller, and E. K. U. Gross, How to represent crystal structures for machine learning: Towards fast prediction of electronic properties, *Phys. Rev. B* **89**, 205118 (2014).
- [7] L. Fiedler, N. A. Modine, S. Schmerler, D. J. Vogel, G. A. Popoola, A. P. Thompson, S. Rajamanickam, and A. Cangi, Predicting electronic structures at any length scale with machine learning, *npj Comput. Mater.* **9**, 115 (2023).
- [8] G. Montavon, M. Rupp, V. Gobre, A. Vazquez-Mayagoitia, K. Hansen, A. Tkatchenko, K.-R. Müller, and O. A. von Lilienfeld, Machine learning of molecular electronic properties in chemical compound space, *New J. Phys.* **15**, 095003 (2013).
- [9] Y. Liu, T. Zhao, W. Ju, and S. Shi, Materials discovery and design using machine learning, *J. Mater.* **3**, 159 (2017).
- [10] D. Lemm, G. F. von Rudorff, and O. A. von Lilienfeld, Machine learning based energy-free structure predictions of molecules, transition states, and solids, *Nat. Commun.* **12**, 4468 (2021).
- [11] A. Merchant, S. Batzner, S. S. Schoenholz, M. Aykol, G. Cheon, and E. D. Cubuk, Scaling deep learning for materials discovery, *Nature (London)* **624**, 80 (2023).
- [12] E. O. Pyzer-Knapp, J. W. Pitera, P. W. J. Staar, S. Takeda, T. Laino, D. P. Sanders, J. Sexton, J. R. Smith, and A. Curioni, Accelerating materials discovery using artificial intelligence, high performance computing and robotics, *npj Comput. Mater.* **8**, 84 (2022).
- [13] V. Stanev, K. Choudhary, A. G. Kusne, J. Paglione, and I. Takeuchi, Artificial intelligence for search and discovery of quantum materials, *Commun. Mater.* **2**, 105 (2021).
- [14] G. Christou, D. Gatteschi, D. N. Hendrickson, and R. Sessoli, Single-molecule magnets, *MRS Bull.* **25**, 66 (2000).
- [15] D. Gatteschi, R. Sessoli, and J. Villain, *Molecular Nanomagnets* (Oxford University Press, Oxford, 2006).
- [16] R. Sessoli, H. L. Tsai, A. R. Schake, S. Wang, J. B. Vincent, K. Folting, D. Gatteschi, G. Christou, and D. N. Hendrickson, High-spin molecules: $[\text{Mn}_{12}\text{O}_{12}(\text{O}_2\text{Cr})_{16}(\text{H}_2\text{O})_4]$, *J. Am. Chem. Soc.* **115**, 1804 (1993).
- [17] P.-H. Lin, T. J. Burchell, L. Ungur, L. F. Chibotaru, W. Wernsdorfer, and M. Murugesu, A polynuclear lanthanide single-molecule magnet with a record anisotropic barrier, *Angew. Chem. Int. Ed.* **48**, 9489 (2009).
- [18] A. Dey, P. Kalita, and V. Chandrasekhar, Lanthanide(III)-based single-ion magnets, *ACS Omega* **3**, 9462 (2018).
- [19] A. Zabala-Lekuona, J. M. Seco, and E. Colacio, Single-molecule magnets: From Mn_{12} -ac to dysprosium metallocenes, a travel in time, *Coord. Chem. Rev.* **441**, 213984 (2021).
- [20] Z. Zhu and J. Tang, Lanthanide single-molecule magnets with high anisotropy barrier: where to from here? *Natl. Sci. Rev.* **9**, nwac194 (2022).
- [21] A. Borah and R. Murugavel, Magnetic relaxation in single-ion magnets formed by less-studied lanthanide ions Ce(III) , Nd(III) , Gd(III) , Ho(III) , Tm(II/III) and Yb(III) , *Coord. Chem. Rev.* **453**, 214288 (2022).

- [22] K. Bernot, C. Daiguebonne, G. Calvez, Y. Suffren, and O. Guillou, A journey in lanthanide coordination chemistry: From evaporable dimers to magnetic materials and luminescent devices, *Acc. Chem. Res.* **54**, 427 (2021).
- [23] G. Gabarró-Riera, G. Aromí, and E. C. Sañudo, Magnetic molecules on surfaces: SMMs and beyond, *Coord. Chem. Rev.* **475**, 214858 (2023).
- [24] L. Bogani and W. Wernsdorfer, Molecular spintronics using single-molecule magnets, *Nat. Mater.* **7**, 179 (2008).
- [25] R. Sessoli, Toward the quantum computer: Magnetic molecules back in the race, *ACS Cent. Sci.* **1**, 473 (2015).
- [26] A. Chiesa, P. Santini, E. Garlatti, F. Luis, and S. Carretta, Molecular nanomagnets: A viable path toward quantum information processing? *Rep. Prog. Phys.* **87**, 034501 (2024).
- [27] A. Gaita-Ariño, F. Luis, S. Hill, and E. Coronado, Molecular spins for quantum computation, *Nat. Chem.* **11**, 301 (2019).
- [28] D. J. Newman, Theory of lanthanide crystal fields, *Adv. Phys.* **20**, 197 (1971).
- [29] A. Abragam and B. Bleaney, *Electron Paramagnetic Resonance of Transition Ions* (Clarendon Press, Oxford, 1970).
- [30] J. J. Moré, The Levenberg-Marquardt algorithm: Implementation and theory, in *Numerical Analysis*, Lecture Notes in Mathematics, edited by G. A. Watson, Vol. 630 (Springer, Berlin, 1978), p. 105–116.
- [31] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. (Cambridge University Press, New York, NY, 1992).
- [32] D. N. Woodruff, R. E. P. Winpenny, and R. A. Layfield, Lanthanide single-molecule magnets, *Chem. Rev. (Washington, DC, U.S.)* **113**, 5110 (2013).
- [33] C. Benelli and D. Gatteschi, *Introduction to Molecular Magnetism* (Wiley, Weinheim 2015).
- [34] O. Waldmann, Relation between electrostatic charge density and spin Hamiltonian models of ligand field in lanthanide complexes, *Inorg. Chem. (Washington, DC, U.S.)* **64**, 1365 (2025).
- [35] B. O. Roos and P.-k. Malmqvist, Relativistic quantum chemistry: The multiconfigurational approach, *Phys. Chem. Chem. Phys.* **6**, 2919 (2004).
- [36] L. Ungur and L. F. Chibotaru, *Ab initio* crystal field for lanthanides, *Chem. Eur. J* **23**, 3708 (2017).
- [37] P. P. Hallmen, C. Köppl, G. Rauhut, H. Stoll, and J. van Slageren, Fast and reliable *ab initio* calculation of crystal field splittings in lanthanide complexes, *J. Chem. Phys.* **147**, 164101 (2017).
- [38] A. Lunghi and F. Totti, The role of anisotropic exchange in single molecule magnets: A CASSCF/NEVPT2 study of the Fe₄ SMM building block [Fe₂(OCH₃)₂(dbm)₄] dimer, *Inorganics* **4**, 28 (2016).
- [39] M. Briganti, G. F. Garcia, J. Jung, R. Sessoli, B. Le Guennic, and F. Totti, Covalency and magnetic anisotropy in lanthanide single molecule magnets: The DyDOTA archetype, *Chem. Sci.* **10**, 7233 (2019).
- [40] J. Jung, M. A. Islam, V. L. Pecoraro, T. Mallah, C. Berthon, and H. Bolvin, Derivation of lanthanide series crystal field parameters from first principles, *Chem. Eur. J* **25**, 15112 (2019).
- [41] X. Wu, J.-F. Li, and B. Yin, The interpretation and prediction of lanthanide single-ion magnets from *ab initio* electronic structure calculation: The capability and limit, *Dalton Trans.* **51**, 14793 (2022).
- [42] A. Soncini and M. Piccardo, *Ab initio* non-covalent crystal field theory for lanthanide complexes: A multiconfigurational non-orthogonal group function approach, *Phys. Chem. Chem. Phys.* **24**, 18915 (2022).
- [43] C. A. Morrison, *Classical Point-Charge Model* (Springer, Berlin, Heidelberg, 1988), Vol. 47.
- [44] C. E. Schäffer and C. K. Jørgensen, The angular overlap model, an attempt to revive the ligand field approaches, *Mol. Phys.* **9**, 401 (1965).
- [45] J. J. Baldoví, A. Gaita-Ariño, and E. Coronado, Modeling the magnetic properties of lanthanide complexes: Relationship of the REC parameters with Pauling electronegativity and coordination number, *Dalton Trans.* **44**, 12535 (2015).
- [46] J. J. Baldoví, J. J. Borrás-Almenar, J. M. Clemente-Juan, E. Coronado, and A. Gaita-Ariño, Modeling the properties of lanthanoid single-ion magnets using an effective point-charge approach, *Dalton Trans.* **41**, 13705 (2012).
- [47] H. Ramanantoanina, W. Urland, F. Cimpoesu, and C. Daul, The angular overlap model extended for two-open-shell *f* and *d* electrons, *Phys. Chem. Chem. Phys.* **16**, 12282 (2014).
- [48] J. J. Baldoví, J. M. Clemente-Juan, E. Coronado, A. Gaita-Ariño, and A. Pali, An updated version of the computational package SIMPRE that uses the standard conventions for Stevens crystal field parameters, *J. Comput. Chem.* **35**, 1930 (2014).
- [49] M. Suta, F. Cimpoesu, and W. Urland, The angular overlap model of ligand field theory for *f* elements: An intuitive approach building bridges between theory and experiment, *Coord. Chem. Rev.* **441**, 213981 (2021).
- [50] J. Mutschler, T. Ruppert, J. Strahinger, S. Schlittenhardt, Z. A. Ali, Y. Peng, C. E. Anson, M. Ruben, A. K. Powell, and O. Waldmann, Transferability of ligand field parameters in a family of 3d-4f M₂Ln₂ butterfly single-molecule magnets, *Inorg. Chem. (Washington, DC, U.S.)* **64**, 6115 (2025).
- [51] A. Lunghi and S. Sanvito, Surfing multiple conformation-property landscapes via machine learning: Designing magnetic anisotropy, *J. Phys. Chem. C* **10**, 5802 (2019).
- [52] Y. Duan, L. E. Rosaleny, J. T. Coutinho, S. Giménez-Santamarina, A. Scheie, J. J. Baldoví, S. Cardona-Serra, and A. Gaita-Ariño, Data-driven design of molecular nanomagnets, *Nat. Commun.* **13**, 7626 (2022).
- [53] Vu Ha Anh Nguyen and A. Lunghi, Predicting tensorial molecular properties with equivariant machine learning models, *Phys. Rev. B* **105**, 165131 (2022).
- [54] V. Zaverkin, J. Netz, F. Zills, A. Köhn, and J. Kästner, Thermally averaged magnetic anisotropy tensors via machine learning based on gaussian moments, *J. Chem. Theory Comput.* **18**, 1 (2022).
- [55] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, Adaptive Computation and Machine Learning (The MIT Press, Cambridge Massachusetts, 2016).
- [56] L. Ardizzone, J. Kruse, S. Wirkert, D. Rahner, E. W. Pellegrini, R. S. Klessen, L. Maier-Hein, C. Rother, and U. Köthe, Analyzing inverse problems with invertible neural networks, in *ICLR* (Vancouver, Canada, 2018).
- [57] Y. Peng, V. Mereacre, C. E. Anson, and A. K. Powell, Butterfly M₂^{III}Er₂ (M^{III} = Fe and Al) SMMs: Synthesis, characterization, and magnetic properties, *ACS Omega* **3**, 6360 (2018).

- [58] G. A. Bain and J. F. Berry, Diamagnetic corrections and Pascal's constants, *J. Chem. Educ.* **85**, 532 (2008).
- [59] D. P. Kingma and M. Welling, An introduction to variational autoencoders, *Found. Trends Mach. Learn.* **12**, 307 (2019).
- [60] I. Higgins, L. Matthey, A. Pal, C. Burgess, X. Glorot, M. Botvinick, S. Mohamed, and A. Lerchner, beta-VAE: Learning basic visual concepts with a constrained variational framework, in *ICLR 2017* (Toulon, France, 2017), <https://openreview.net/forum?id=Sy2fzU9gl>.
- [61] L. Dinh, J. Sohl-Dickstein, and S. Bengio, Density estimation using real NVP, in *ICLR 2017* (Toulon, France, 2017).
- [62] A. Gretton, K. M. Borgwardt, M. J. Rasch, B. Schölkopf, and A. Smola, A kernel two-sample test, *J. Mach. Learn. Res.* **13**, 723 (2012).
- [63] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/7tfd-22jr> for Figs. S1, S2, S3, S4, S5 and Tables S1, S2.
- [64] W. Boulila, M. Driss, E. Alshantiti, M. Al-Sarem, F. Saeed, and M. Krichen, Weight initialization techniques for deep learning algorithms in remote sensing: Recent trends and future perspectives, in *Advances on Smart and Soft Computing*, edited by F. Saeed, T. Al-Hadhrani, E. Mohammed, and M. Al-Sarem (Springer Singapore, Singapore, 2022), pp. 477–484.
- [65] N. Altman and M. Krzywinski, The curse(s) of dimensionality, *Nat. Methods* **15**, 399 (2018).
- [66] A. Beskos, M. Girolami, S. Lan, P. E. Farrell, and A. M. Stuart, Geometric MCMC for infinite-dimensional inverse problems, *J. Comput. Phys.* **335**, 327 (2017).
- [67] Edited by C. Sammut and G. I. Webb, *Encyclopedia of Machine Learning* (Springer, New York and London, 2010).
- [68] Y. Chen, L. Lu, G. E. Karniadakis, and L. Dal Negro, Physics-informed neural networks for inverse problems in nano-optics and metamaterials, *Opt. Express* **28**, 11618 (2020).
- [69] M. Raissi, P. Perdikaris, and G. E. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, *J. Comput. Phys.* **378**, 686 (2019).