Rapid Detection of Strong Correlation with Machine Learning for Transition-Metal Complex High-Throughput Screening

Fang Liu[1], Chenru Duan[1,2], and Heather J. Kulik[1,]*

[1]*Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA, 02139*
[2]*Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA, 02139*

AUTHOR INFORMATION

Corresponding Author

*email: hjkulik@mit.edu, phone: 617-253-4584

ABSTRACT: Despite its widespread use in chemical discovery, approximate density functional theory (DFT) is poorly suited to many targets, such as those containing open-shell, $3d$ transition metals that can be expected to have strong multi-reference (MR) character. For discovery workflows to be predictive, we need automated, low-cost methods that can distinguish the regions of chemical space where DFT should be applied from those where it should not. We curate over 4,800 open-shell transition-metal complexes up to hundreds of atoms in size from prior high-throughput DFT studies and evaluate affordable, finite-temperature DFT evaluation of fractional occupation number (FON)-based MR diagnostics. We show that intuitive measures of strong correlation (i.e., the HOMO–LUMO gap) are not predictive of MR character as judged by FON-based diagnostics. Analysis of independently trained machine learning (ML) models to predict HOMO–LUMO gaps and FON-based diagnostics reveals differences in metal- and ligand-sensitivity of the two quantities. We use our trained ML models to rapidly evaluate MR character over a space of ca. 187,000 theoretical complexes, identifying large-scale trends in spin-state-dependent MR character and finding small HOMO–LUMO gap complexes while ensuring low MR character.

TOC GRAPHICS
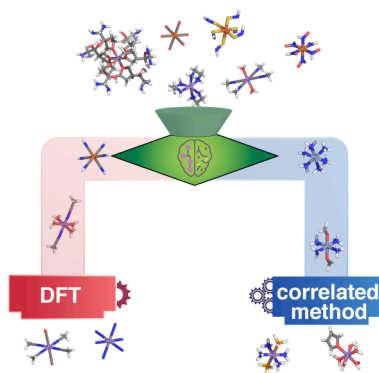
High-throughput computational screening is essential in the discovery of new molecules[1-5] and materials[6-9]. For these efforts, density functional theory (DFT) with approximate exchange-correlation functionals is nearly exclusively employed to ensure the low computational cost and ease of automation needed for the study of thousands of molecules with tens to hundreds of atoms. Open-shell transition-metal complexes that are useful catalysts[4, 10-16] and functional materials[17-21] exemplify both the promise and outstanding challenges for computational screening. The properties of these molecules can be tuned by variation of ligand chemistry, which in turn influences the relative stability of spin- and oxidation- states[5, 22], highlighting the opportunities for large-scale screening in their design[5, 23-24].

In small-scale studies, careful benchmarking of the computational method is possible[25-28], including the detection of strong correlation that motivates the use of multireference (MR) electronic structure techniques.[29-31] However, high-scaling methods have been limited in large-scale screening contexts to small organic molecules.[32-33] The development of machine learning (ML) property prediction models[23, 34-38] and interatomic potentials[39-43] compounds the challenges associated with screening, both by motivating the generation of large datasets and by magnifying the impact of electronic structure biases when they are learned by an ML model.

First-row open-shell transition-metal complexes with nearly degenerate $3d$ orbitals can be expected to exhibit strong correlation.[25-27, 44-45] Here, approximate DFT will fail[46-49], and the accurate treatment of dynamic and static correlation in larger transition-metal complexes remains an active area of research[29-31]. For small, well-studied molecules (e.g., the chromium dimer[50-53]), the effects of strong correlation are understood, but target complexes in chemical discovery lack literature precedent for guiding expectations of MR character. To advance large-scale screening in these spaces, it is necessary to develop tools to detect strong correlation that can be integrated

into existing DFT workflows at no greater complexity or computational cost.

While numerous MR diagnostics have been developed[54-65], we focus on diagnostics[63-65] based on fractional occupation numbers (FONs) that can be obtained from computationally affordable finite-temperature DFT[66] (FT-DFT). Since electron correlation can be attributed alternately to both dynamical and non-dynamical (i.e., MR) contributions[67-69], Matito and coworkers[63-64] have derived expressions for dynamical, I_D, and non-dynamical, I_{ND}, quantities. The I_D is computed from the occupation, n, of orbital i with spin σ as[63-64]:

$$I_D = \frac{1}{4}\sum_{\sigma,i}\left[n_i^\sigma\left(1-n_i^\sigma\right)\right]^{1/2} - \frac{1}{2}\sum_{\sigma,i}n_i^\sigma\left(1-n_i^\sigma\right) \tag{1}$$

The non-dynamical[63-64], I_{ND}, is:

$$I_{ND} = \frac{1}{2}\sum_{\sigma,i}n_i^\sigma\left(1-n_i^\sigma\right) \tag{2}$$

where larger I_{ND} values can be used as an indicator of MR character[33, 63-64, 70]. The ratio,[67] r_{ND}, of the I_{ND} to the total correlation (i.e., I_D+I_{ND}) has been proposed by Martin and coworkers as a size-intensive MR diagnostic:

$$r_{ND} = \frac{I_{ND}}{I_D + I_{ND}} \tag{3}$$

Although I_D and I_{ND} are not strictly bounded, larger values of r_{ND} are indicative of strong MR character.[33, 67]

Grimme and coworkers have developed the closely related fractional occupation density (FOD)[65], which can be integrated over all space to yield N_{FOD}, a size-extensive FON-based diagnostic for quantifying non-dynamical correlation. We focus on the size-intensive r_{ND}, and we demonstrate our approach is general both to size-extensive, FON-based diagnostics (i.e., I_{ND}, N_{FOD}) and approximately size-intensive quantities that we obtain by normalizing with the number of valence electrons, n_{ve} (i.e., I_{ND}/n_{ve} and N_{FOD}/n_{ve}).

While FON-based diagnostics have demonstrated promise[65, 70] for the quantitative

prediction of MR character in studies of small sets of transition-metal complexes, they have yet to be incorporated in large-scale, high-throughput screening. We thus first curated a diverse set of 4,865 mononuclear octahedral transition-metal complexes from six prior studies[71-76] and assessed their MR character with FON-based diagnostics. We accelerated the calculation of FON-based diagnostics by leveraging the existing converged geometries and wavefunctions from fixed-occupation DFT for the FT-DFT[66] calculations (see Computational Details). All complexes contain a mid-row transition metal in one of two oxidation states (i.e., M(II)/M(III) where M = Cr, Mn, Fe, or Co) in its high-spin (HS) or low-spin (LS) state (see Computational Details). Over all sets considered, the metals and spin states are evenly distributed, with only somewhat fewer Cr complexes (Figure 1).
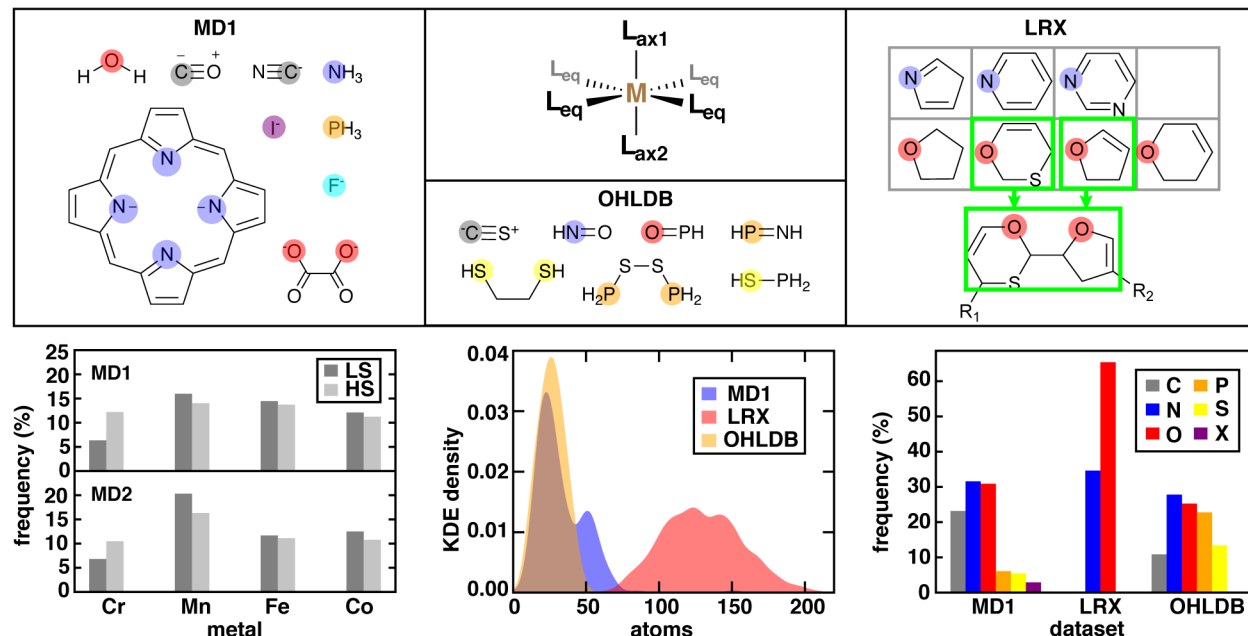

Figure 1. Properties of the octahedral complexes in the datasets studied in this work (*MD1, MD2,* and the two component datasets, *LRX* and *OHLDB*, included in *MD2* only). (top) Schematic of an a transition metal complex illustrating possible unique (one equatorial, L_{eq}, and up to two axial, L_{ax1} and L_{ax2}) and representative ligands. Ligand atoms that coordinate the metal are shaded with circles colored by element: O in red, N in blue, P in orange, S in yellow, C in gray, F in cyan, and I in purple. (bottom, left) Bar graph of the metal identity and spin state (HS: dark gray, LS: light gray) of *MD1* and *MD2*. (bottom, center) Kernel density estimation (KDE) of the distribution of the size of complexes in *MD1, LRX,* and *OHLDB*. (bottom, right) Clustered bar graph for the connecting atom identity (X indicates any halide) in *MD1, LRX,* and *OHLDB*.

A 2,305-complex subset (*MD1*) from four of the studies[71-74] consists of small to mid-sized complexes constructed from common ligands originally to study their spin[71-72, 74] or redox[73] properties (Figure 1 and Supporting Information Table S1). The *MD1* complexes are around 50 atoms in size and primarily contain first-row (i.e., C, N, or O) coordinating atoms (Figure 1 and Supporting Information Table S1). The larger 4,865-complex *MD2* dataset consists of all complexes in *MD1* along with: i) complexes of heuristically enumerated ligands (*OHLDB*[76]) and ii) large bidentate redox (*LRX*[75]) complexes (Figure 1 and Supporting Information Tables S1–S2). The *OHLDB* complexes are smaller (ca. 25 atoms) and have more diverse coordinating atoms and bonding than in *MD1*, whereas the *LRX* complexes are much larger (i.e., up to 200 atoms) but with fewer coordinating atom types and more uniform bonding (Figure 1 and Supporting Information Table S1). The diversity of size and bonding in *MD2* means that we are less likely to have information from prior experiment or computation in comparison to the more common complexes in *MD1*.

Over the *MD1* dataset, the r_{ND} values span a large range (0.10–0.66), suggesting a significant variation in the extent of strong correlation, although all complexes are derived from common ligands (Figure 2). The *MD2* dataset has a comparable r_{ND} range, despite its greater diversity in both ligand chemistry and complex size (Supporting Information Figure S1). The size-extensive diagnostics (i.e., I_{ND} and N_{FOD}) differ significantly between *MD1* and *MD2*, but the comparable behavior of r_{ND} as well as n_{ve}-normalized quantities (i.e., I_{ND}/n_{ve} and N_{FOD}/n_{ve}) support our focus on a size-insensitive diagnostic (Supporting Information Figures S2–S5).
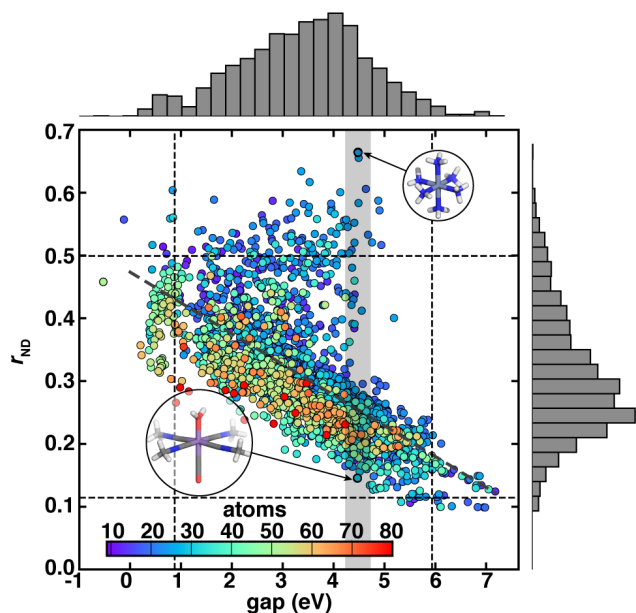
Figure 2. Scatter plot of the 2D distribution of HOMO–LUMO gap (eV) and r_{ND} values for *MD1* complexes. The symbols are colored by the number of atoms in the complex, as indicated in inset colorbar. 1D histograms are aligned with the respective axes. Vertical and horizontal dashed lines indicate two standard deviations around the mean. The correlation between r_{ND} and gap is indicated with a gray dashed line (R^2=0.413). A gray shaded region (HOMO–LUMO gap = 4.48 ± 0.25 eV) is shown with the complexes with the highest (LS $Cr^{II}(NH_3)_6$) and lowest (HS $Mn^{II}(misc)_4(H_2O)(CO)$) r_{ND} values outlined and shown with inset structures.

One might expect it to be straightforward to predict MR character from fixed-occupation DFT calculations, e.g., by the size of the gap between the highest-occupied and lowest-unoccupied molecular orbitals (HOMO–LUMO gap). The HOMO–LUMO gaps span a wide range (0–7 eV) over the *MD1* set, but they only have a weak correlation (R^2=0.41) with the r_{ND} diagnostic (Figure 2). Still, the relationship follows expectations, with the smallest HOMO–LUMO gap complexes having more MR character (i.e., higher r_{ND}) than the complexes with the largest HOMO–LUMO gaps. Nevertheless, this relationship worsens over the diverse *MD2* set (R^2=0.17), despite HOMO–LUMO gap and r_{ND} ranges comparable to *MD1* (Supporting Information Figure S1). We might expect this to be due to greater size-dependence of the HOMO–LUMO gap than r_{ND}, but correlations for both sets are still weak ($R^2 < 0.6$) with size-extensive (i.e., I_{ND}, N_{FOD}) MR diagnostics (Figure 2 and Supporting Information Figures S2, S4,

7

and S6 and Table S3).

Given the weak relationship between the HOMO–LUMO gap and MR character, we investigated what gives rise to low or high MR character in *MD1* for small (i.e., 0.88 ± 0.25 eV) and large (i.e., 4.48 ± 0.25 eV) gap values. Many of the 276 large-gap complexes have expectedly low r_{ND}, including a HS $Mn^{II}(misc)_4(H_2O)(CO)$ with an r_{ND} value (0.13) in the bottom 2% for *MD1* (Figure 2). Conversely, LS $Cr^{II}(NH_3)_6$ has the highest r_{ND} (0.66) in all of *MD1*, despite its large HOMO–LUMO gap, suggesting that LS Cr complexes might have especially high MR character (Figure 2). For the 66 small HOMO–LUMO gap *MD1* complexes, both the expected high r_{ND} values are observed (i.e., LS $Mn^{III}(H_2O)_5(furan)$, r_{ND} = 0.60) but unexpectedly low MR character is as well (i.e., LS $Mn^{III}(pisc)_6$, r_{ND} = 0.27, Figure 2). Given the higher diversity of *MD2*, distinct deviations are observed, including low r_{ND} and gaps (e.g., in an *LRX* O-coordinating bidentate homoleptic HS Mn^{III} complex, r_{ND} = 0.22 and gap = 0.3 eV) even for complexes with similar metal and coordinating atoms to much higher r_{ND} *MD1* points (Supporting Information Figure S1). Despite the intuitive link between MR character and HOMO–LUMO gap, metal identity (e.g., Cr), or spin state (e.g., LS), exceptions are apparent, motivating the development of ML models capable of encoding these complex relationships.

We trained ML (i.e., artificial neural network, ANN, and kernel ridge regression, KRR) models to independently predict FON-based diagnostics (e.g., r_{ND}) and frontier orbital energetics (i.e., the HOMO level and HOMO–LUMO gap, see Computational Details and Supporting Information Text S1). We used revised autocorrelation (RAC)[77-78] functions as input features, which have previously been demonstrated for predictive, geometry-free estimation of frontier orbital energies to good (ca. 0.1–0.2 eV) accuracy (Supporting Information Text S2).[73, 77] RACs are products and differences on the molecular graph of heuristic properties (i.e., topology,

identity, nuclear charge, covalent radius, or electronegativity) of pairs of atoms a fixed number, d, of bonds apart (see Computational Details and Supporting Information Text S2). For the ML models trained on the *MD2* dataset with RAC features, ANN model test mean absolute errors (MAEs, HOMO: 0.26 eV and HOMO–LUMO gap: 0.31 eV) are similar to those for *MD1* or prior work[10, 14] on *MD1* subsets, increasing only slightly despite being obtained on a much more diverse dataset (Figure 3 and Supporting Information Table S4 and Figures S7–S8). As shown in prior work[75, 79], the greater diversity in size and ligand chemistry of *MD2* complexes means a model trained only on *MD1* would not be expected to be predictive on *MD2*.
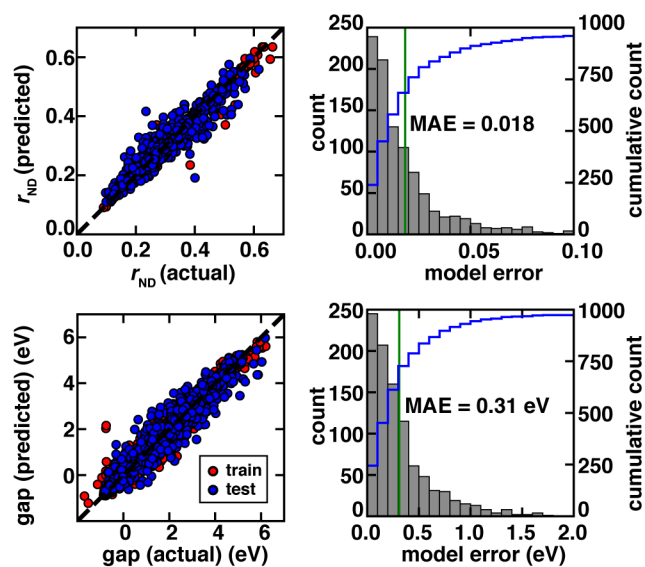

Figure 3. ANN model performance for predicting r_{ND} (upper) and HOMO–LUMO gap (lower) for the *MD2* data set. (left) Parity plots of actual vs predicted values for train (red filled circles) and test (blue filled circles) data points along with a black dotted parity line. (right) Distribution of absolute test set model errors for r_{ND} (unitless, bins: 0.005) and HOMO–LUMO gap (in eV, bins: 0.1 eV) with the MAE annotated as green vertical bars and the cumulative count shown in blue according to the axis on the right.

It is not evident *a priori* if FON-based diagnostics that are derived from temperature-dependent properties obtained with FT-DFT are as easily learned quantities as the fixed-occupation DFT frontier orbital energies. The ML model trained on *MD2* to predict r_{ND} does exhibit as good performance as that for the HOMO–LUMO gap, with the ANN test MAE for r_{ND}

(0.018) corresponds to a smaller mean absolute percent error (3%) over the r_{ND} range than for the HOMO–LUMO gap (4%, Figure 3). The r_{ND} MAE represents a modest error in comparison to the variation between diagnostic values associated with low (e.g., $r_{ND} < 0.3$) or high (e.g., $r_{ND} > 0.5$) MR character (Figure 3 and Supporting Information Table S4 and Figure S7). In comparison to predicting r_{ND} with a linear mapping from the HOMO–LUMO gap, the ANN performance is vastly superior (linear: $R^2 = 0.456$ vs ANN: $R^2 = 0.942$) and test MAEs are significantly lower (Supporting Information Figures S9–S10 and Table S5).

Interpretation of the character of informative RACs provides insight[77, 80] into why MR character (i.e., r_{ND}) and HOMO–LUMO gap differ among these transition-metal complexes. To identify the most informative features[77, 80], we pair feature selection with the training of KRR models[73, 77] with similar performance to that obtained with ANNs (Supporting Information Tables S4, S6–S12, and Figures S11–S12). Analysis of feature sets selected by training on *MD1* confirms that distinct properties are important for predicting r_{ND} (i.e., r_{ND}-28) versus predicting the HOMO–LUMO gap (gap-28) or HOMO level (Figure 4 and Supporting Information Figure S13 and Tables S7–S8, S13–S15). Although the r_{ND}-28 and gap-28 feature sets are the same size, the gap subset contains a higher portion of global features (43%) than the r_{ND} subset (32%, Figure 4 and Supporting Information Tables S14–S15). Electronegativity-based RACs are more crucial for predicting MR character (25% of r_{ND}-28 vs 11% of gap-28) as is the metal nuclear charge, whereas global, covalent-radius-based features present only in gap-28 highlight the greater size-dependence of the HOMO–LUMO gap (Figure 4 and Supporting Information Tables S14–S15). Qualitatively similar trends are observed for models trained on *MD2*, although the size of the gap-selected feature set grows significantly due to greater diversity of ligand chemistry in the set (Supporting Information Tables S10–S11).
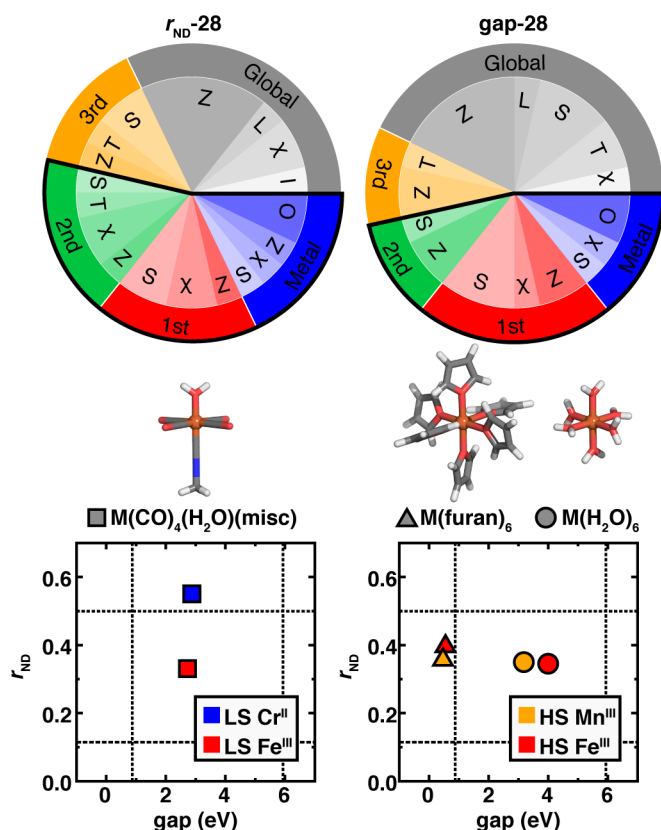
Figure 4. (upper) Pie charts of the features selected by random-forest-ranked recursive-feature addition (RF-RFA) on the *MD1* data set for HOMO–LUMO gap (gap-28, left) and the r_{ND} diagnostic (r_{ND}-28, right). Features are grouped by the most metal-distant atoms: metal in blue, first coordination sphere in red, second coordination sphere in green, third coordination sphere in orange, or more distant, global features in gray (a black outline is used to group the first three categories). Within each distance category, the property (i.e., χ, S, T, Z, or I) is also indicated, with oxidation/spin state (O) assigned as metal-local and ligand charge (L) assigned as global. (lower) HOMO–LUMO gap (eV) and r_{ND} values (unitless) illustrating the larger effect of metal identity on r_{ND} (left) and of non-local features on gap (right). Vertical and horizontal dashed lines indicate two standard deviations around the mean gap or r_{ND} for *MD1*. The representative complex structures are shown along with their symbol legends above the plots.

Although over our datasets we had observed a weak correlation between the gap and MR character, these feature-selected subsets provide explicit design principles over *MD1*, i.e., to separately tailor the gap and MR character by focusing ligand-size-based versus metal-based properties. For example, the LS $Fe^{III}(CO)_4(H_2O)$(misc) and $Cr^{II}(CO)_4(H_2O)$(misc) complexes differ only in their metal center and have comparable HOMO–LUMO gaps (Fe(III): 2.73 eV, Cr(II) 2.87 eV, Figure 4). This difference instead has a profound effect on MR character, with

the Cr^{II} complex having a high r_{ND} (0.55), whereas the Fe^{III} complex r_{ND} (0.33) is lower (Figure 4). Conversely, replacing water ligands with larger furan ligands in HS homoleptic Fe^{III} or Mn^{III} complexes has a modest effect on MR character (r_{ND} from 0.34 to 0.40 in Fe^{III}) because the immediate coordinating environment is unchanged, whereas the increased ligand size reduces the gap substantially (Fe^{III}: by 3.44 eV, Figure 4).

To demonstrate the value of our approach in a context representative of chemical discovery efforts, we applied the gap and r_{ND} ANN models trained on *MD2* to a space of 187,200 transition metal complexes. This space of theoretical complexes contains HS and LS M(II/III) (M = Cr, Mn, Fe, or Co) centers in complex with 36 unique ligands derived from the original *MD1* dataset. Although all ligand identities were in the ANN training data, only 1% (1,836) of the complexes were, and thus the theoretical complexes have distinct properties (e.g., size and charge) from the original dataset (Supporting Information Tables S16–S18 and Figures S14–S15).

Over this large theoretical complex space, we investigate what the ML model has learned about how MR character is influenced by transition-metal complex chemistry (i.e., spin state, metal, and ligand). Spin-state ordering calculations in particular are known to be sensitive to method choice[80-86], especially due to the lack of error cancellation when the degree of correlation recovery differs between spin states[30, 86-89]. It has been observed, for example, in small-scale studies of organic molecules[68, 90] and some first-row transition metal complexes[86] that LS states are more degenerate and thus have higher MR character than HS states.

Although some metal centers in our large-scale set follow this trend, others do not. The Cr and Mn complexes exhibit an increase in MR character from HS to LS for both r_{ND} and other FON-based diagnostics, whereas Fe and Co complexes have limited spin-state dependence

(Figure 5 and Supporting Information Figures S16–S21). For the HS complexes in either oxidation state, all metals have comparable diagnostic (e.g., r_{ND}) distributions, suggesting the HS–LS MR character difference for Cr/Mn arises primarily from increased LS MR character (Figure 5 and Supporting Information Figures S17–S21). Indeed, for the Cr/Mn complexes with the greatest HS–LS MR character difference, the LS r_{ND} is very high while the HS r_{ND} is somewhat lower than the overall average (Supporting Information Figure S22). The ligands that most frequently appear in these complexes are small, sometimes charged species (e.g. NH_3, OH^-, and S^{2-}, CN^-) with intermediate field strengths (Supporting Information Figure S17). Although MR character is strongly spin-state- and metal-dependent, subtle differences in chemistry are apparent between ligands that frequently occur in low-MR character complexes (e.g., NH_2CH_3) and those (e.g., NCO^-) in complexes with high MR character (Supporting Information Figures S23–S24).
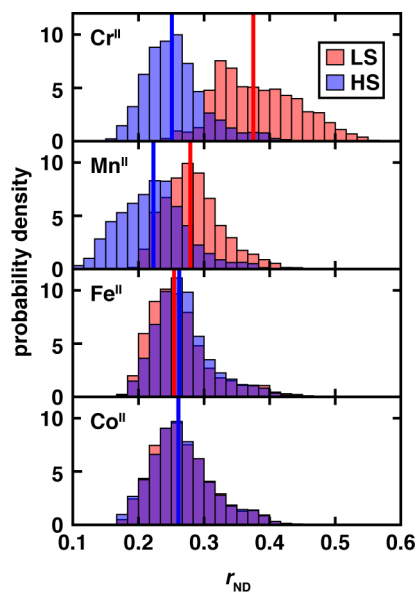

Figure 5. Normalized probability density distribution of r_{ND} (unitless, bin size 0.0167) as predicted by the ANN for the 93600 M(II) complexes (M=Cr, Mn, Fe, Co) in the full compound space. There are 11700 complexes for each M(II) spin state. The histograms are colored by spin (red for low spin, LS, and blue for high spin, HS). The median of each distribution is indicated by a vertical line with the same corresponding color.

Within the space of theoretical complexes, independent ML models also make it possible

to target optimal DFT properties (e.g., for the HOMO–LUMO gap) while identifying DFT-safe portions of chemical space by avoiding complexes with high MR character. ML-model-predicted MR character and HOMO–LUMO gap vary widely but are, as in the training data, weakly correlated (Supporting Information Figures S25–S29). These trends are expected to be robust because ANN model test errors are modest, but we focus our analysis on low uncertainty points in the theoretical space as quantified by those with the smallest distances in ANN model latent space[79] (Supporting Information Text S3 and Figures S30–S31).

Given the expectation[68, 91] that small HOMO–LUMO gap complexes will have strong MR character, we sought to identify chemical motifs that break this relationship (Figure 6). For the complexes with confidently-assigned small HOMO–LUMO gaps (< 1 eV) and below-average MR character (r_{ND} < 0.3), clear patterns emerge (Figure 6). One in three (134 of 403) of these target complexes are from the same family of HS Co^{III} or Mn^{III} complexes with equatorial substituted-pyridine ligands and axial weak-field ligands that thus reside in a narrow region of the complex space (Figure 6 and Supporting Information Tables S19–S20 and Figures S32–S33). Conversely, nearby complexes (e.g., with strong-field equatorial carbonyl ligands) also have small gaps but instead have relatively high MR character (Figure 6 and Supporting Information Figure S34 and Table S21). In the future, these ANN models could be paired with optimization strategies[72, 75, 92-93] for the continued discovery of DFT-safe islands of complexes with optimal properties along with low MR character and high ML model confidence from even larger spaces of theoretical transition-metal complexes.
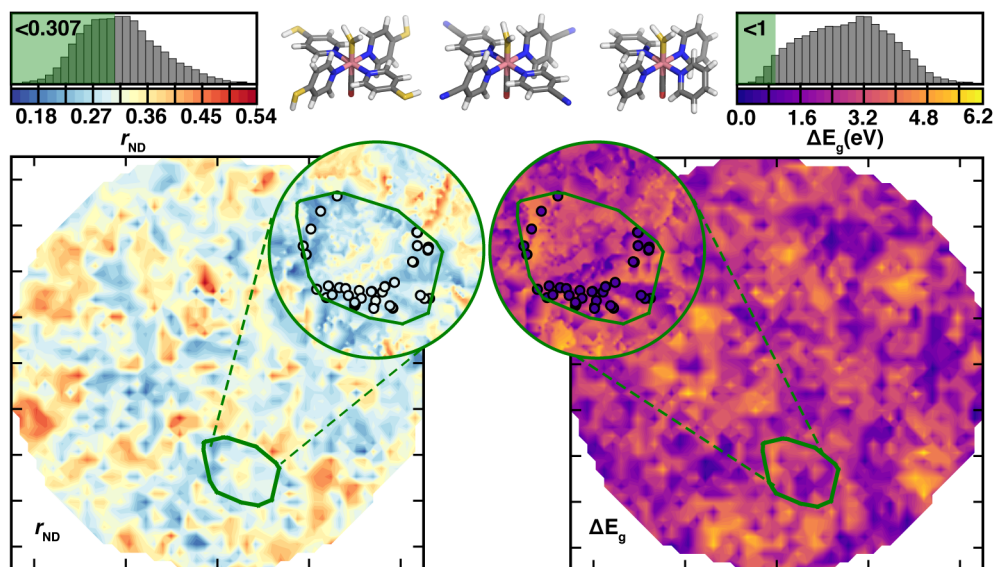
Figure 6. Properties of the theoretical complex space colored by ANN predicted r_{ND} (left, unitless) and HOMO–LUMO gap (right, ΔE_g in eV). (top) The 1D histograms of r_{ND} and HOMO–LUMO gap are aligned with their respective color bars and target zones (r_{ND} <0.307, HOMO–LUMO gap < 1 eV) shaded in green. (bottom) The t-distributed stochastic neighbor embedding plots are shown and colored as indicated in top inset color bar. The convex hull of a family of 3744 octahedral complexes with functionalized pyridinyl ligands is indicated by a green outline. Inset circles show a zoom of this convex hull with discrete complexes in circles (representative complexes shown in top middle of plot): the 134 complexes within target ranges of r_{ND} and HOMO–LUMO gap and low (bottom 10%) distance in latent space to training data.

In summary, we have demonstrated a low-cost approach that can be integrated into current DFT high-throughput screening workflows for open-shell transition-metal complexes and materials. Over a curated set of nearly 5,000 open-shell transition-metal complexes from prior high-throughput DFT studies, we evaluated MR character from FON-based diagnostics. Because intuitive measures of strong correlation (i.e., the HOMO–LUMO gap) were not predictive of these FON-based diagnostics, we trained ML models to separately predict frontier orbital energies (i.e., the HOMO–LUMO gap) and FON-based diagnostics. Models trained to predict finite-temperature-derived FON-based diagnostics were as predictive as models trained to predict fixed-occupation DFT properties. Analysis of the important features for model prediction revealed that MR character is more metal-sensitive, whereas the HOMO–LUMO gap is more ligand-sensitive. We used these principles to identify opportunities for tailoring the two

15

quantities independently. We demonstrated the promise of our ANN models for chemical discovery efforts by evaluating MR character over a space of over 187,000 theoretical complexes, identifying large-scale trends in spin-state-dependent MR character, and discovering small HOMO–LUMO gap complexes with low MR character. We expect this approach to be valuable both for the practical goal of identifying where DFT-level workflows are sufficiently robust and in discovering strongly correlated molecules as test cases for more advanced electronic structure methods.

Computational Details

Datasets and calculations. We curated a dataset of 4,865 mononuclear octahedral transition-metal complexes generated in six prior studies[71-76] using fixed-occupation DFT geometry optimizations in TeraChem[94-95] automated with molSimplify[96-97]. All sets used the same high-spin (HS) and low-spin (LS) multiplicity definitions in mid-row M(III)/M(II) complexes as follows: quintet-singlet for both d^4 Mn(III)/Cr(II) and d^6 Co(III)/Fe(II), sextet-doublet for d^5 Fe(III)/Mn(II), and quartet-doublet for both d^3 Cr(III) and d^7 Co(II).

We automated the calculation of FT-DFT[66] MR diagnostics[63-65] (e.g., r_{ND} and N_{FOD}) with MultirefPredict[67] and the QCEngine[98] interface to TeraChem[94-95] on the curated structures from prior work. For consistency, all DFT calculations employed the B3LYP[99-101] hybrid functional with LANL2DZ[102] effective core potentials for the transition metals, I, or Br and the 6-31G* basis for the remaining atoms. Level-shifting[103] was used in unrestricted (i.e., non-singlet) calculations with a uniform 0.25 a.u. value for FT-DFT and most prior fixed-occupation DFT calculations (Supporting Information Table S22). The FT-DFT calculations employed the recommended[104] temperature for B3LYP (9000 K) with Fermi–Dirac smearing and were initiated from fixed-occupation DFT wavefunctions[71-76] when available (Supporting Information

Text S4 and Tables S22–S23).

All complexes were filtered prior to analysis and ML model training in a four-step process. Complexes for which FT-DFT calculations did not converge were first eliminated followed by those with positive (i.e., unbound) or unreported majority-spin HOMO levels from fixed-occupation DFT (Supporting Information Text S4 and Table S24). Geometries were next checked[73] for preserved connectivity and quality, and, lastly, highly symmetry-broken fixed-occupation DFT (i.e., with $<S^2>$ deviations from $S(S+1) > 1$) results were excluded (Supporting Information Table S25 and Table S24).

ML models. KRR and ANN models were trained on all 151 RACs (i.e., $42d+30$ with a cutoff of $d = 3$ after elimination of invariant RACs) along with three overall (i.e., oxidation state, denticity and total charge)[77] complex features as well as feature-selected subsets (Supporting Information Text S2). Hyperparameter optimization for all models was carried out with Hyperopt[105] using a random 80% train/20% test split, with 20% of the training set (16% overall) set aside as the validation subset for hyperparameter selection. Input features and outputs were normalized over the training set to have zero mean and unit variance. As in prior work[73], KRR model feature selection was carried out with random forest[106]-ranked recursive feature addition (RF-RFA)[107]. With each feature addition, the KRR model was trained (i.e., to select regularization strength and kernel width) and judged on the KRR model R^2 for the validation set. RF-RFA was stopped when no improvement of R^2 was observed. Final KRR models were implemented in scikit-learn[108], trained on the full (80% overall) training set, and tested on the 20% test set (Supporting Information Table S26).

Fully connected ANN models were trained using Keras[109] with TensorFlow[110] as the backend and Hyperopt[105] for hyperparameter selection (Supporting Information Table S27). Each

ANN model was trained with batch optimization for 1000 epochs with early stopping when no

more performance improvement was observed, and we included dropout regularization[111-112] to

avoid overfitting (Supporting Information Table S28). The optimal ANN topology for all models

consisted of 512 nodes per layer with either two (e.g., for the HOMO level) or three (e.g., for

HOMO−LUMO gap and r_{ND}) three hidden layers (Supporting Information Table S28).

ASSOCIATED CONTENT

Supporting Information. The following files are available free of charge.

Statistics of denticity and size and net charge of complexes in data sets; correlation between gap
and r_{ND} for *MD2*; correlation between gap and MR diagnostics (r_{ND}, I_{ND}, I_{ND}/n_{ve} N_{FOD}, N_{FOD}/n_{ve})
for *MD1* and *MD2*; KDE of r_{ND} and gap for molecules of different sizes; regression parameters
for molecules of different sizes; definition of HOMO and LUMO; extended description of RAC
featurization; performance of ML models for HOMO-LUMO gap and r_{ND}; train and test set
performance of KRR and ANN models trained on *MD1* and *MD2*; *MD1*; linear model for
predicting r_{ND} from DFT or ANN predicted gap; additional linear, KRR, and ANN model
performance statistics; performance of ML models for HOMO level, I_{ND}, and N_{FOD}; optimal
KRR features for HOMO prediction and gap prediction of *MD1* and *MD2*; pie chart of the
optimal features for HOMO level prediction of *MD1*; statistics on HOMO-52, gap-28, and r_{ND} -
28 feature sets; ligands used for building the theoretical complex space; allowed ligand
combinations and spin/oxidation combinations in the theoretical complex space; distribution of
sizes in the theoretical complex space; distribution of theoretical space M(III) r_{ND} values;
difference of theoretical space HS and LS r_{ND} values; distribution of theoretical space
M(II)/M(III) I_{ND}, I_{ND}/n_{ve}, N_{FOD}, and N_{FOD}/n_{ve} values; contributions to high HS-LS r_{ND} difference
complexes; ligand types in the top 5% and bottom 5% of r_{ND} in the theoretical complex space;
correlation between gap and MR diagnostics (r_{ND}, I_{ND}, N_{FOD}, I_{ND}/n_{ve}, N_{FOD}/n_{ve}) for 187k
theoretical complexes; details of latent space distance evaluation; distribution of gap latent space
distances and r_{ND} latent space distances over theoretical space; ligands in the target leads inside
the convex hull of theoretical space; metals in the target leads inside the convex hull of
theoretical space; the r_{ND} distribution and HOMO-LUMO distribution of target leads inside the
convex hull; convex hull compounds with large r_{ND}; attributes of ligands in convex hull
compounds with large r_{ND}; summary of size of data sets after each stage of refinement;
additional details about finite-temperature calculations; number of core orbitals for different
elements; summary of wavefunction related information for each dataset; geometry check cutoffs
for data refinement; optimal hyperparameters for KRR models in this work; hyperspace for ANN
hyperparameter optimization; optimal hyperparameters for ANN models in this work. (PDF)

Total energies in fixed-occupation and finite-temperature DFT, MR diagnostics, and orbital energies in fixed occupation-DFT of all molecules in the six raw datasets and the refined datasets *MD1* and *MD2*; list of molecules eliminated during *MD2* set curation; list of molecules in *MD2* lacking fixed-occupation DFT wavefunction information from previous studies; ANN predicted MR diagnostics and orbital energies for the design space; list of unique ligands in *MD1* and *MD2*; the ANN and KRR models trained on *MD1* and *MD2* training set; geometries of all molecules in *MD2* (ZIP)

AUTHOR INFORMATION

Notes

The authors declare no competing financial interests.

 REFERENCES
(1)   Shu, Y.; Levine, B. G. Simulated Evolution of Fluorophores for Light Emitting Diodes. *J. Chem. Phys.* **2015,** *142*, 104104.
(2)   Gomez-Bombarelli, R.; Aguilera-Iparraguirre, J.; Hirzel, T. D.; Duvenaud, D.; Maclaurin, D.; Blood-Forsythe, M. A.; Chae, H. S.; Einzinger, M.; Ha, D. G.; Wu, T.; et al.

Design of Efficient Molecular Organic Light-Emitting Diodes by a High-Throughput Virtual Screening and Experimental Approach. *Nat. Mater.* **2016,** *15*, 1120-+.

(3)   Kanal, I. Y.; Owens, S. G.; Bechtel, J. S.; Hutchison, G. R. Efficient Computational Screening of Organic Polymer Photovoltaics. *J. Phys. Chem. Lett.* **2013,** *4*, 1613-1623.

(4)   Vogiatzis, K. D.; Polynski, M. V.; Kirkland, J. K.; Townsend, J.; Hashemi, A.; Liu, C.; Pidko, E. A. Computational Approach to Molecular Catalysis by 3d Transition Metals: Challenges and Opportunities. *Chem. Rev.* **2018,** *119*, 2453-2523.

(5)   Foscato, M.; Jensen, V. R. Automated in Silico Design of Homogeneous Catalysts. *ACS Catal.* **2020,** *10*, 2354-2377.

(6)   Curtarolo, S.; Hart, G. L.; Nardelli, M. B.; Mingo, N.; Sanvito, S.; Levy, O. The High-Throughput Highway to Computational Materials Design. *Nat. Mater.* **2013,** *12*, 191-201.

(7)   Ong, S. P.; Richards, W. D.; Jain, A.; Hautier, G.; Kocher, M.; Cholia, S.; Gunter, D.; Chevrier, V. L.; Persson, K. A.; Ceder, G. Python Materials Genomics (Pymatgen): A Robust, Open-Source Python Library for Materials Analysis. *Comput. Mater. Sci.* **2013,** *68*, 314-319.

(8)   Nørskov, J. K.; Bligaard, T. The Catalyst Genome. *Angew. Chem., Int. Ed.* **2013,** *52*, 776-777.

(9)   Meredig, B.; Agrawal, A.; Kirklin, S.; Saal, J. E.; Doak, J.; Thompson, A.; Zhang, K.; Choudhary, A.; Wolverton, C. Combinatorial Screening for New Materials in Unconstrained Composition Space with Machine Learning. *Phys. Rev. B* **2014,** *89*, 094104.

(10)   Huo, P.; Uyeda, C.; Goodpaster, J. D.; Peters, J. C.; Miller III, T. F. Breaking the Correlation between Energy Costs and Kinetic Barriers in Hydrogen Evolution via a Cobalt Pyridine-Diimine-Dioxime Catalyst. *ACS Catal.* **2016,** *6*, 6114-6123.

(11)   Grajciar, L.; Heard, C. J.; Bondarenko, A. A.; Polynski, M. V.; Meeprasert, J.; Pidko, E. A.; Nachtigall, P. Towards Operando Computational Modeling in Heterogeneous Catalysis. *Chem. Soc. Rev.* **2018,** *47*, 8307-8348.

(12)   Arockiam, P. B.; Bruneau, C.; Dixneuf, P. H. Ruthenium(II)-Catalyzed C-H Bond Activation and Functionalization. *Chem. Rev.* **2012,** *112*, 5879-5918.

(13)   Schultz, D. M.; Yoon, T. P. Solar Synthesis: Prospects in Visible Light Photocatalysis. *Science* **2014,** *343*, 985-+.

(14)   Shaffer, D. W.; Bhowmick, I.; Rheingold, A. L.; Tsay, C.; Livesay, B. N.; Shores, M. P.; Yang, J. Y. Spin-State Diversity in a Series of Co(Ii) PNP Pincer Bromide Complexes. *Dalton Trans.* **2016,** *45*, 17910-17917.

(15)   Tsay, C.; Yang, J. Y. Electrocatalytic Hydrogen Evolution under Acidic Aqueous Conditions and Mechanistic Studies of a Highly Stable Molecular Catalyst. *J. Am. Chem. Soc.* **2016,** *138*, 14174-14177.

(16)   Schilling, M.; Patzke, G. R.; Hutter, J.; Luber, S. Computational Investigation and Design of Cobalt Aqua Complexes for Homogeneous Water Oxidation. *J. Phys. Chem. C* **2016,** *120*, 7966-7975.

(17)   Ashley, D. C.; Jakubikova, E. Ironing out the Photochemical and Spin-Crossover Behavior of Fe (II) Coordination Compounds with Computational Chemistry. *Coord. Chem. Rev.* **2017,** *337*, 97-111.

(18)   Bowman, D. N.; Bondarev, A.; Mukherjee, S.; Jakubikova, E. Tuning the Electronic Structure of Fe(II) Polypyridines via Donor Atom and Ligand Scaffold Modifications: A Computational Study. *Inorg. Chem.* **2015,** *54*, 8786-8793.

(19)   Yella, A.; Lee, H. W.; Tsao, H. N.; Yi, C. Y.; Chandiran, A. K.; Nazeeruddin, M. K.; Diau, E. W. G.; Yeh, C. Y.; Zakeeruddin, S. M.; Gratzel, M. Porphyrin-Sensitized Solar Cells

with Cobalt (II/III)-Based Redox Electrolyte Exceed 12 Percent Efficiency. *Science* **2011,** *334*, 629-634.

(20)   Kuttipillai, P. S.; Zhao, Y. M.; Traverse, C. J.; Staples, R. J.; Levine, B. G.; Lunt, R. R. Phosphorescent Nanocluster Light-Emitting Diodes. *Adv. Mater.* **2016,** *28*, 320-326.

(21)   Linfoot, C. L.; Leitl, M. J.; Richardson, P.; Rausch, A. F.; Chepelin, O.; White, F. J.; Yersin, H.; Robertson, N. Thermally Activated Delayed Fluorescence (TADF) and Enhancing Photoluminescence Quantum Yields of Cu-I(Diimine)(Diphosphine)(+) Complexes-Photophysical, Structural, and Computational Studies. *Inorg. Chem.* **2014,** *53*, 10854-10861.

(22)   Swart, M.; Costas, M. *Spin States in Biochemistry and Inorganic Chemistry: Influence on Structure and Reactivity*. John Wiley & Sons: 2015.

(23)   Meyer, B.; Sawatlon, B.; Heinen, S.; von Lilienfeld, O. A.; Corminboeuf, C. Machine Learning Meets Volcano Plots: Computational Discovery of Cross-Coupling Catalysts. *Chem. Sci.* **2018,** *9*, 7069-7077.

(24)   Weymuth, T.; Reiher, M. Inverse Quantum Chemistry: Concepts and Strategies for Rational Compound Design. *Int. J. Quantum Chem.* **2014,** *114*, 823-837.

(25)   Jiang, W.; DeYonker, N. J.; Wilson, A. K. Multireference Character for 3d Transition-Metal-Containing Molecules. *J. Chem. Theory Comput.* **2012,** *8*, 460-468.

(26)   Wang, J.; Manivasagam, S.; Wilson, A. K. Multireference Character for 4d Transition Metal-Containing Molecules. *J. Chem. Theory Comput.* **2015,** *11*, 5865-5872.

(27)   DeYonker, N. J.; Peterson, K. A.; Steyl, G.; Wilson, A. K.; Cundari, T. R. Quantitative Computational Thermochemistry of Transition Metal Species. *J. Phys. Chem. A* **2007,** *111*, 11269-11277.

(28)   Zhang, W.; Truhlar, D. G.; Tang, M. Tests of Exchange-Correlation Functional Approximations against Reliable Experimental Data for Average Bond Energies of 3d Transition Metal Compounds. *J. Chem. Theory Comput.* **2013,** *9*, 3965-3977.

(29)   Gagliardi, L.; Truhlar, D. G.; Li Manni, G.; Carlson, R. K.; Hoyer, C. E.; Bao, J. L. Multiconfiguration Pair-Density Functional Theory: A New Way to Treat Strongly Correlated Systems. *Acc. Chem. Res.* **2017,** *50*, 66-73.

(30)   Phung, Q. M.; Feldt, M.; Harvey, J. N.; Pierloot, K. Toward Highly Accurate Spin State Energetics in First-Row Transition Metal Complexes: A Combined CASPT2/Cc Approach. *J. Chem. Theory Comput.* **2018,** *14*, 2446-2455.

(31)   Singh, S. K.; Eng, J.; Atanasov, M.; Neese, F. Covalency and Chemical Bonding in Transition Metal Complexes: An Ab Initio Based Ligand Field Perspective. *Coord. Chem. Rev.* **2017,** *344*, 2-25.

(32)   Smith, J. S.; Nebgen, B. T.; Zubatyuk, R.; Lubbers, N.; Devereux, C.; Barros, K.; Tretiak, S.; Isayev, O.; Roitberg, A. E. Approaching Coupled Cluster Accuracy with a General-Purpose Neural Network Potential through Transfer Learning. *Nat. Commun.* **2019,** *10*, 1-8.

(33)   Duan, C.; Liu, F.; Nandy, A.; Kulik, H. J. Data-Driven Approaches Can Overcome the Cost-Accuracy Tradeoff in Multireference Diagnostics. *J. Chem. Theory Comput.* **2020,** *16*, 4373-4387.

(34)   Collins, C. R.; Gordon, G. J.; von Lilienfeld, O. A.; Yaron, D. J. Constant Size Descriptors for Accurate Machine Learning Models of Molecular Properties. *J. Chem. Phys.* **2018,** *148*, 241718.

(35)   Bleiziffer, P.; Schaller, K.; Riniker, S. Machine Learning of Partial Charges Derived from High-Quality Quantum-Mechanical Calculations. *J. Chem. Inf. Model.* **2018,** *58*, 579-590.

(36)   De, S.; Bartok, A. P.; Csanyi, G.; Ceriotti, M. Comparing Molecules and Solids across

Structural and Alchemical Space. *Phys. Chem. Chem. Phys.* **2016,** *18*, 13754-13769.

(37)   Ward, L.; Agrawal, A.; Choudhary, A.; Wolverton, C. A General-Purpose Machine Learning Framework for Predicting Properties of Inorganic Materials. *npj Comput. Mater.* **2016,** *2*, 16028.

(38)   Pilania, G.; Wang, C.; Jiang, X.; Rajasekaran, S.; Ramprasad, R. Accelerating Materials Property Predictions Using Machine Learning. *Sci. Rep.* **2013,** *3*, 2810.

(39)   Yao, K.; Herr, J. E.; Toth, D. W.; Mckintyre, R.; Parkhill, J. The Tensormol-0.1 Model Chemistry: A Neural Network Augmented with Long-Range Physics. *Chem. Sci.* **2018,** *9*, 2261-2269.

(40)   Behler, J. Perspective: Machine Learning Potentials for Atomistic Simulations. *J. Chem. Phys.* **2016,** *145*, 170901.

(41)   Smith, J. S.; Isayev, O.; Roitberg, A. E. ANI-1: An Extensible Neural Network Potential with DFT Accuracy at Force Field Computational Cost. *Chem. Sci.* **2017,** *8*, 3192-3203.

(42)   Zhang, L.; Han, J.; Wang, H.; Car, R.; Weinan, E. Deep Potential Molecular Dynamics: A Scalable Model with the Accuracy of Quantum Mechanics. *Phys. Rev. Lett.* **2018,** *120*, 143001.

(43)   Chmiela, S.; Tkatchenko, A.; Sauceda, H. E.; Poltavsky, I.; Schütt, K. T.; Müller, K.-R. Machine Learning of Accurate Energy-Conserving Molecular Force Fields. *Sci. Adv.* **2017,** *3*, e1603015.

(44)   Gaggioli, C. A.; Stoneburner, S. J.; Cramer, C. J.; Gagliardi, L. Beyond Density Functional Theory: The Multiconfigurational Approach to Model Heterogeneous Catalysis. *ACS Catal.* **2019,** *9*, 8481-8502.

(45)   Boguslawski, K.; Tecmer, P.; Legeza, O.; Reiher, M. Entanglement Measures for Single- and Multireference Correlation Effects. *J. Phys. Chem. Lett.* **2012,** *3*, 3129-3135.

(46)   Yu, H. S.; Li, S. L.; Truhlar, D. G. Perspective: Kohn-Sham Density Functional Theory Descending a Staircase. *J. Chem. Phys.* **2016,** *145*, 130901.

(47)   Cohen, A. J.; Mori-Sánchez, P.; Yang, W. Challenges for Density Functional Theory. *Chem. Rev.* **2012,** *112*, 289-320.

(48)   Becke, A. D. Perspective: Fifty Years of Density-Functional Theory in Chemical Physics. *J. Chem. Phys.* **2014,** *140*, 18A301.

(49)   Cramer, C. J.; Truhlar, D. G. Density Functional Theory for Transition Metals and Transition Metal Chemistry. *Phys. Chem. Chem. Phys.* **2009,** *11*, 10757-10816.

(50)   Guo, S.; Watson, M. A.; Hu, W.; Sun, Q.; Chan, G. K.-L. N-Electron Valence State Perturbation Theory Based on a Density Matrix Renormalization Group Reference Function, with Applications to the Chromium Dimer and a Trimer Model of Poly (P-Phenylenevinylene). *J. Chem. Theory Comput.* **2016,** *12*, 1583-1591.

(51)   Müller, T. Large-Scale Parallel Uncontracted Multireference-Averaged Quadratic Coupled Cluster: The Ground State of the Chromium Dimer Revisited. *J. Phys. Chem. A* **2009,** *113*, 12729-12740.

(52)   Pantazis, D. A. Meeting the Challenge of Magnetic Coupling in a Triply-Bridged Chromium Dimer: Complementary Broken-Symmetry Density Functional Theory and Multireference Density Matrix Renormalization Group Perspectives. *J. Chem. Theory Comput.* **2019,** *15*, 938-948.

(53)   Goodgame, M. M.; Goddard III, W. A. The" Sextuple" Bond of Chromium Dimer. *J. Phys. Chem.* **1981,** *85*, 215-217.

(54)   Lee, T. J.; Taylor, P. R. A Diagnostic for Determining the Quality of Single-Reference

Electron Correlation Methods. *Int. J. Quantum Chem.* **1989**, 199-207.

(55)   Sears, J. S.; Sherrill, C. D. Assessing the Performance of Density Functional Theory for the Electronic Structure of Metal-Salens: The d(2)-Metals. *J. Phys. Chem. A* **2008,** *112*, 6741-6752.

(56)   Sears, J. S.; Sherrill, C. D. Assessing the Performance of Density Functional Theory for the Electronic Structure of Metal-Salens: The 3d(0)-Metals. *J. Phys. Chem. A* **2008,** *112*, 3466-3477.

(57)   Langhoff, S. R.; Davidson, E. R. Configuration Interaction Calculations on the Nitrogen Molecule. *Int. J. Quantum Chem.* **1974,** *8*, 61-72.

(58)   Janssen, C. L.; Nielsen, I. M. B. New Diagnostics for Coupled-Cluster and Moller-Plesset Perturbation Theory. *Chem Phys Lett* **1998,** *290*, 423-430.

(59)   Nielsen, I. M. B.; Janssen, C. L. Double-Substitution-Based Diagnostics for Coupled-Cluster and Moller-Plesset Perturbation Theory. *Chem Phys Lett* **1999,** *310*, 568-576.

(60)   Fogueri, U. R.; Kozuch, S.; Karton, A.; Martin, J. M. L. A Simple DFT-Based Diagnostic for Nondynamical Correlation. *Theor. Chem. Acc.* **2013,** *132*, 1291.

(61)   Tishchenko, O.; Zheng, J. J.; Truhlar, D. G. Multireference Model Chemistries for Thermochemical Kinetics. *J. Chem. Theory Comput.* **2008,** *4*, 1208-1219.

(62)   Schultz, N. E.; Zhao, Y.; Truhlar, D. G. Density Functionals for Inorganometallic and Organometallic Chemistry. *J. Phys. Chem. A* **2005,** *109*, 11127-11143.

(63)   Ramos-Cordoba, E.; Salvador, P.; Matito, E. Separation of Dynamic and Nondynamic Correlation. *Phys. Chem. Chem. Phys.* **2016,** *18*, 24015-24023.

(64)   Ramos-Cordoba, E.; Matito, E. Local Descriptors of Dynamic and Nondynamic Correlation. *J. Chem. Theory Comput.* **2017,** *13*, 2705-2711.

(65)   Grimme, S.; Hansen, A. A Practicable Real-Space Measure and Visualization of Static Electron-Correlation Effects. *Angew. Chem., Int. Ed.* **2015,** *54*, 12308-12313.

(66)   Weinert, M.; Davenport, J. W. Fractional Occupations and Density-Functional Energies and Forces. *Phys. Rev. B* **1992,** *45*, 13709-13712.

(67)   Kesharwani, M. K.; Sylvetsky, N.; Köhn, A.; Tew, D. P.; Martin, J. M. L. Do CCSD and Approximate CCSD-F12 Variants Converge to the Same Basis Set Limits? The Case of Atomization Energies. *J. Chem. Phys.* **2018,** *149*, 154109.

(68)   Sprague, M. K.; Irikura, K. K. Quantitative Estimation of Uncertainties from Wavefunction Diagnostics. In *Thom H. Dunning, Jr.*; Springer: 2015; pp 307-318.

(69)   Via-Nadal, M.; Rodríguez-Mayorga, M.; Ramos-Cordoba, E.; Matito, E. Singling out Dynamic and Nondynamic Correlation. *J. Phys. Chem. Lett.* **2019,** *10*, 4032-4037.

(70)   Süß, D.; Huber, S. E.; Mauracher, A. On the Impact of Multi-Reference Character of Small Transition Metal Compounds on Their Bond Dissociation Energies. *J. Chem. Phys.* **2020,** *152*, 114104.

(71)   Janet, J. P.; Kulik, H. J. Predicting Electronic Structure Properties of Transition Metal Complexes with Neural Networks. *Chem. Sci.* **2017,** *8*, 5137-5152.

(72)   Janet, J. P.; Chan, L.; Kulik, H. J. Accelerating Chemical Discovery with Machine Learning: Simulated Evolution of Spin Crossover Complexes with an Artificial Neural Network. *J. Phys. Chem. Lett.* **2018,** *9*, 1064-1071.

(73)   Nandy, A.; Duan, C.; Janet, J. P.; Gugler, S.; Kulik, H. J. Strategies and Software for Machine Learning Accelerated Discovery in Transition Metal Chemistry. *Ind. Eng. Chem. Res.* **2018,** *57*, 13973-13986.

(74)   Duan, C.; Janet, J. P.; Liu, F.; Nandy, A.; Kulik, H. J. Learning from Failure: Predicting

Electronic Structure Calculation Outcomes with Machine Learning Models. *J. Chem. Theory Comput.* **2019,** *15*, 2331-2345.

(75)   Janet, J. P.; Ramesh, S.; Duan, C.; Kulik, H. J. Accurate Multiobjective Design in a Space of Millions of Transition Metal Complexes with Neural-Network-Driven Efficient Global Optimization. *ACS Cent. Sci.* **2020,** *6*, 513-524.

(76)   Gugler, S.; Janet, J. P.; Kulik, H. J. Enumeration of De Novo Inorganic Complexes for Chemical Discovery and Machine Learning. *Mol. Sys. Des. Eng.* **2020,** *5*, 139-152.

(77)   Janet, J. P.; Kulik, H. J. Resolving Transition Metal Chemical Space: Feature Selection for Machine Learning and Structure-Property Relationships. *J. Phys. Chem. A* **2017,** *121*, 8939-8954.

(78)   Janet, J. P.; Gani, T. Z. H.; Steeves, A. H.; Ioannidis, E. I.; Kulik, H. J. Leveraging Cheminformatics Strategies for Inorganic Discovery: Application to Redox Potential Design. *Ind. Eng. Chem. Res.* **2017,** *56*, 4898-4910.

(79)   Janet, J. P.; Duan, C.; Yang, T.; Nandy, A.; Kulik, H. J. A Quantitative Uncertainty Metric Controls Error in Neural Network-Driven Chemical Discovery. *Chem. Sci.* **2019,** *10*, 7913-7922.

(80)   Janet, J. P.; Liu, F.; Nandy, A.; Duan, C.; Yang, T.; Lin, S.; Kulik, H. J. Designing in the Face of Uncertainty: Exploiting Electronic Structure and Machine Learning Models for Discovery in Inorganic Chemistry. *Inorg. Chem.* **2019,** *58*, 10592-10606.

(81)   Ioannidis, E. I.; Kulik, H. J. Towards Quantifying the Role of Exact Exchange in Predictions of Transition Metal Complex Properties. *J. Chem. Phys.* **2015,** *143*, 034104.

(82)   Ganzenmüller, G.; Berkaïne, N.; Fouqueau, A.; Casida, M. E.; Reiher, M. Comparison of Density Functionals for Differences between the High-(T 2 G 5) and Low-(a 1 G 1) Spin States of Iron (II) Compounds. IV. Results for the Ferrous Complexes [Fe (L)('Nhs 4')]. *J. Chem. Phys.* **2005,** *122*, 234321.

(83)   Droghetti, A.; Alfè, D.; Sanvito, S. Assessment of Density Functional Theory for Iron (II) Molecules across the Spin-Crossover Transition. *J. Chem. Phys.* **2012,** *137*, 124303.

(84)   Mortensen, S. R.; Kepp, K. P. Spin Propensities of Octahedral Complexes from Density Functional Theory. *J. Phys. Chem. A* **2015,** *119*, 4041-4050.

(85)   Pierloot, K. The CASPT2 Method in Inorganic Electronic Spectroscopy: From Ionic Transition Metal to Covalent Actinide Complexes∗. *Mol. Phys.* **2003,** *101*, 2083-2094.

(86)   Pierloot, K.; Phung, Q. M.; Domingo, A. Spin State Energetics in First-Row Transition Metal Complexes: Contribution of (3s3p) Correlation and Its Description by Second-Order Perturbation Theory. *J. Chem. Theory Comput.* **2017,** *13*, 537-553.

(87)   Flöser, B. M.; Guo, Y.; Riplinger, C.; Tuczek, F.; Neese, F. Detailed Pair Natural Orbital-Based Coupled Cluster Studies of Spin Crossover Energetics. *J. Chem. Theory Comput.* **2020,** *16*, 2224-2235.

(88)   Feldt, M.; Phung, Q. M.; Pierloot, K.; Mata, R. A.; Harvey, J. N. Limits of Coupled-Cluster Calculations for Non-Heme Iron Complexes. *J. Chem. Theory Comput.* **2019,** *15*, 922-937.

(89)   Zhang, D.; Truhlar, D. G. Spin Splitting Energy of Transition Metals: A New, More Affordable Wave Function Benchmark Method and Its Use to Test Density Functional Theory. *J. Chem. Theory Comput.* **2020,** *16*, 4416-4428.

(90)   Mizukami, W.; Kurashige, Y.; Yanai, T. Communication: Novel Quantum States of Electron Spins in Polycarbenes from Ab Initio Density Matrix Renormalization Group Calculations. *J. Chem. Phys.* **2010,** *133*, 091101.

(91) Shen, J.; Xu, E.; Kou, Z.; Li, S. New Coupled Cluster Approaches Based on the Unrestricted Hartree–Fock Reference for Treating Molecules with Multireference Character. *Phys. Chem. Chem. Phys.* **2011,** *13*, 8795-8804.

(92) Herbol, H. C.; Hu, W.; Frazier, P.; Clancy, P.; Poloczek, M. Efficient Search of Compositional Space for Hybrid Organic–Inorganic Perovskites via Bayesian Optimization. *npj Comput. Mater.* **2018,** *4*, 51.

(93) Häse, F.; Roch, L. M.; Kreisbeck, C.; Aspuru-Guzik, A. Phoenics: A Bayesian Optimizer for Chemistry. *ACS Cent. Sci.* **2018,** *4*, 1134-1145.

(94) Petachem, L. Petachem. http://www.petachem.com (accessed July 19, 2020).

(95) Ufimtsev, I. S.; Martinez, T. J. Quantum Chemistry on Graphical Processing Units. 3. Analytical Energy Gradients, Geometry Optimization, and First Principles Molecular Dynamics. *J. Chem. Theory Comput.* **2009,** *5*, 2619-2628.

(96) Ioannidis, E. I.; Gani, T. Z. H.; Kulik, H. J. molSimplify: A Toolkit for Automating Discovery in Inorganic Chemistry. *J. Comput. Chem.* **2016,** *37*, 2106-2117.

(97) KulikGroup. molSimplify Documentation. http://molsimplify.mit.edu (accessed July 19, 2020).

(98) MolSSI. QCEngine. https://github.com/MolSSI/QCEngine (accessed July 19, 2020).

(99) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **1994,** *98*, 11623-11627.

(100) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993,** *98*, 5648-5652.

(101) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988,** *37*, 785-789.

(102) Hay, P. J.; Wadt, W. R. Ab Initio Effective Core Potentials for Molecular Calculations. Potentials for the Transition Metal Atoms Sc to Hg. *J. Chem. Phys.* **1985,** *82*, 270-283.

(103) Saunders, V. R.; Hillier, I. H. A "Level–Shifting" Method for Converging Closed Shell Hartree–Fock Wave Functions. *Int. J. Quantum Chem.* **1973,** *7*, 699-705.

(104) Liu, F.; Duan, C.; Kulik, H. J. Multirefpredict. https://github.com/hjkgrp/MultirefPredict (accessed July 19, 2020).

(105) Bergstra, J.; Yamins, D.; Cox, D. D. In *Hyperopt: A Python Library for Optimizing the Hyperparameters of Machine Learning Algorithms*, Proceedings of the 12th Python in science conference, Year; pp 13-20.

(106) Breiman, L. Random Forests. *Mach. Learn.* **2001,** *45*, 5-32.

(107) Guyon, I.; Elisseeff, A. An Introduction to Variable and Feature Selection. *J. Mach. Learn. Res.* **2003,** *3*, 1157-1182.

(108) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V. Scikit-Learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011,** *12*, 2825-2830.

(109) Chollet, F. Keras. https://keras.io (accessed July 19, 2020).

(110) Abadi, M.; Agarwal, A.; Barham, P.; Brevdo, E.; Chen, Z.; Citro, C.; Corrado, G. S.; Davis, A.; Dean, J.; Devin, M.; et al. *Tensorflow: Large-Scale Machine Learning on Heterogeneous Systems*, 2015.

(111) Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *J. Mach. Learn. Res.* **2014,** *15*, 1929-1958.

(112)   Hinton, G. E.; Srivastava, N.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. R. Improving Neural Networks by Preventing Co-Adaptation of Feature Detectors. *arXiv preprint arXiv:1207.0580* **2012**.