Chemical Science

ROYAL SOCIETY OF CHEMISTRY

EDGE ARTICLE

View Article Online
View Journal | View Issue

Cite this: *Chem. Sci.*, 2023, **14**, 1419

All publication charges for this article have been paid for by the Royal Society of Chemistry

Low-cost machine learning prediction of excited state properties of iridium-centered phosphors†

Gianmarco G. Terrones, [a] Chenru Duan, [ab] Aditya Nandy [ab] and Heather J. Kulik *[ab]

Received 7th November 2022
Accepted 5th January 2023

DOI: 10.1039/d2sc06150c

rsc.li/chemical-science

Prediction of the excited state properties of photoactive iridium complexes challenges *ab initio* methods such as time-dependent density functional theory (TDDFT) both from the perspective of accuracy and of computational cost, complicating high-throughput virtual screening (HTVS). We instead leverage low-cost machine learning (ML) models and experimental data for 1380 iridium complexes to perform these prediction tasks. We find the best-performing and most transferable models to be those trained on electronic structure features from low-cost density functional tight binding calculations. Using artificial neural network (ANN) models, we predict the mean emission energy of phosphorescence, the excited state lifetime, and the emission spectral integral for iridium complexes with accuracy competitive with or superseding that of TDDFT. We conduct feature importance analysis to determine that high cyclometalating ligand ionization potential correlates to high mean emission energy, while high ancillary ligand ionization potential correlates to low lifetime and low spectral integral. As a demonstration of how our ML models can be used for HTVS and the acceleration of chemical discovery, we curate a set of novel hypothetical iridium complexes and use uncertainty-controlled predictions to identify promising ligands for the design of new phosphors while retaining confidence in the quality of the ANN predictions.

[a]*Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

[b]*Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

† Electronic supplementary information (ESI) available: Structures of **CN** and **NN** ligands from the experimental dataset; histograms of the target properties in the experimental dataset; information about DFT calculations; information about the CSD iridium complex search; information about the feature sets and their features; comparison between different fingerprint similarity metrics and charge schemes; correlation of xTB features with themselves, with target properties, and with DFT features; performance of different feature sets on the random and grouped split; the ranking by MAE of different feature sets on the random and grouped split; the change in model performance from random to grouped split for different feature sets; the effect of UQ cutoff on model accuracy in predicting for lifetime and spectral integral; comparison between different ML models; distribution of xTB features; effect of ligand substitution on lifetime and spectral integral; ligands present in complexes with extreme predicted properties; list of complexes used for TDDFT benchmarking; correlation coefficients between experiment, ANN predictions, and TDDFT predictions; comparison of experiment, ANN predictions, and TDDFT predictions for $Em_{50/50}$ and lifetime; ANN predictions for complexes with long experimental lifetime; confusion matrices for a 2 μs lifetime cutoff; impact of a 70/30 train test split; most dissimilar HLS ligands as determined by Dice similarity; hyperparameters of the best-performing ANNs; attrition of CSD complexes; comparison of singlet and triplet geometries of phosphors; CAM-B3LYP and ωB97X-D3BJ TDDFT energy predictions (PDF). XYZ files of the **CN** and **NN** ligands from the experimental dataset; XYZ files of ligands mined from the CSD; xTB features of experimental and hypothetical phosphors; ANN-predicted values for the hypothetical phosphors; train/validation/test splits for the random and grouped splits; example Python scripts for featurization, ANN training, and ANN application; example files for ORCA TDDFT calculation (ZIP). See DOI: https://doi.org/10.1039/d2sc06150c

1. Introduction

Interactions between light and matter underpin phenomena ranging from photovoltaics[1] to photosynthesis[2] to bioluminescence,[3] and the design of functional materials that can leverage these interactions has led to significant technological advancements.[4–6] Exemplary of these advancements are photoactive iridium complexes that have been investigated extensively due to their applications in lighting and display technology,[7–10] photocatalysis,[11–13] and bioimaging.[14,15] The spin–orbit coupling (SOC) characteristic of iridium causes these complexes to efficiently convert excitons into light or chemical energy.[16] Simultaneously, iridium uniquely limits nonradiative decay rates by destabilizing a metal-centered (^{3}MC) triplet excited state due to strong metal–ligand bonding,[17] further improving efficiency. In iridium-centered complexes, the judicious selection of ligands allows for the modulation of phosphorescence color (*i.e.*, emission wavelength) and efficiency/brightness by modulating excited state lifetime and photoluminescence quantum yield.

The desired excited state properties in these highly tunable phosphors are application-dependent. In these complexes, emission energies span the visible spectrum, with complexes at the extremes of the distribution emitting red (1.6 eV)[18] or blue light (2.8 eV).[19] Furthermore, excited state lifetimes in these complexes are on the scale of microseconds, with shorter lifetimes (under 2 μs)[16] preferred for displays and longer lifetimes

preferred for photocatalysis[11] and bioimaging.[14] In addition, for display technologies and bioimaging a high photoluminescence quantum yield is desired. The accurate prediction of these excited state properties will enable the discovery of novel iridium complexes for vibrant display technologies and green photocatalysis.

To screen a large number of compounds, computational modeling with time-dependent density functional theory (TDDFT) can be used for affordable predictions of some properties of transition metal complexes. While TDDFT methods are commonly employed to estimate emission energies,[20–28] the calculation of lifetimes and quantum yields is more challenging both from an accuracy and a computational cost perspective. The calculation of lifetime[23,29–33] requires the inclusion of SOC in TDDFT to estimate the transition dipole moment between the sublevels of the excited triplet (*i.e.*, T_1) and the ground state (*i.e.*, S_0). The calculation of photoluminescence quantum yield further requires the calculation of nonradiative rates, which entails the use of thermal vibration correlation function rate theory[34–36] and excited state geometry optimization.[36,37] Thus, while *ab initio* computational methods have provided valuable insight into the properties of iridium complexes, they are computation-intensive, requiring around one day of computation time per complex for the least-demanding calculations, and may not reach the accuracy required to enable rational design.

Supervised machine learning (ML) has emerged as a powerful complement to *ab initio* methods in recent years due to its capacity to reproduce *ab initio* results at significantly lower cost,[38–42] enabling the screening of vast regions of chemical space.[43] Furthermore, ML models can be trained on experimental data, enabling the prediction of properties that challenge *ab initio* methods, such as material stability.[44] With regard to excited state properties, ML models have been successfully applied for the prediction of phosphorescence energies,[45] fluorescence rates,[46] and fluorescence energies and quantum yields[47] after training on *ab initio* or experimental data. While ML models were first demonstrated for accelerating DFT screening of Ir catalysts in 2020,[48] the extension to directly predicting experimental catalytic[49] or photophysical[50,51] properties has only recently been demonstrated. The need to predict and optimize multiple properties of iridium complexes that challenge TDDFT motivates the continued extension of ML to the direct prediction of experimental properties.

In this work, we use ML to predict three key properties of iridium complexes: $Em_{50/50}$ (mean emission energy), excited state lifetime, and emission spectral integral (brightness). We train and evaluate artificial neural networks (ANNs) on a recent experimental dataset[52] of 1380 iridium(III) phosphors and their properties. This large experimental dataset represents an ideal scenario for ML model training given its uniformity in comparison to acquiring heterogeneous data from multiple sources and conditions. We show that features generated with density functional tight binding lead to the most predictive ANN performance and generalization on out-of-sample data. Using these features, we identify trends in phosphor properties, and we extend our models to a new set of hypothetical iridium phosphors. These experimentally-informed ANNs enable fast, accurate prediction of iridium phosphor properties for the rapid exploration of chemical space when paired with uncertainty control to only apply the ANNs where they are likely to be predictive.

2. Data and representations

2.1. Dataset

We built the structures of bidentate ligands used in the experimental study of DiLuzio *et al.*[52] on Ir(III) complexes of the form

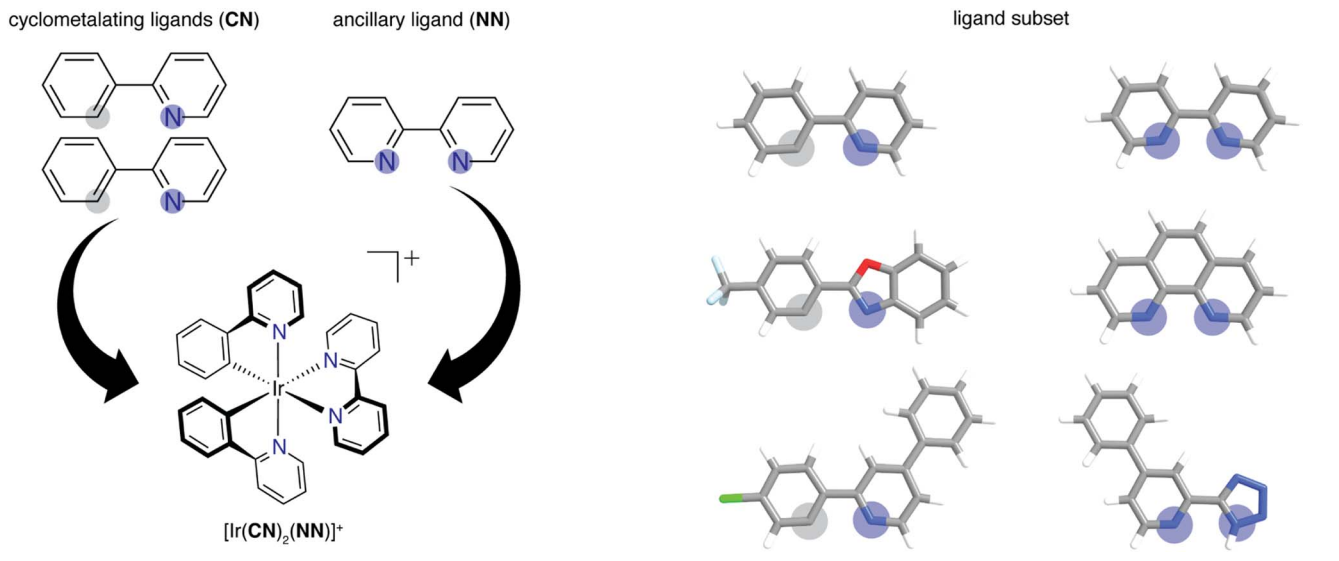

Fig. 1 (Left) Schematic of how two identical CN ligands and one NN ligand comprise each of the iridium phosphors studied in this work. Coordinated nitrogen (carbon) atoms are indicated with blue (gray) circles. (Right) Examples of CN and NN ligands in the experimental dataset of 1380 iridium phosphors. Atoms are colored as follows: white for hydrogen, gray for carbon, blue for nitrogen, red for oxygen, light blue for fluorine, and green for chlorine.

$[Ir(\mathbf{CN})_2(\mathbf{NN})]^+$ (Fig. 1). We assigned each ligand as either cyclometalating (**CN**) or ancillary (**NN**), as determined by the two iridium-coordinating atom identities. We studied the same 60 **CN** ligands and 23 **NN** ligands from the prior experimental study,[52] excluding only the monodentate DMSO ligand in the prior work, giving rise to a combinatorial set of 1380 $[Ir(\mathbf{CN})_2(\mathbf{NN})]^+$ phosphor complexes. This set of 83 ligands will be referred to as the high-throughput ligand set (HLS), and we use the same labeling as in the prior study when referring to individual ligands (ESI Tables S1 and S2†). We used experimental data from the prior study[52] on the three target properties, $Em_{50/50}$, excited state lifetime, and emission spectral integral. The experimental values for these properties were reported for each of the 1380 iridium phosphors in DMSO solvent and were used for ML model training and performance assessment (ESI Fig. S1†). **CN** ligands were generated in their neutral form (*i.e.*, with a proton added) for featurization. Because of this, all ligands are neutral with the exception of three **NN** ligands (ESI Text S1†). After ligand construction using the draw tool in Avogadro v1.1.2 (ref. 53 and 54) and force field (*i.e.*, UFF) optimization, we used these ligands to generate the structures of all possible iridium complexes with one distinct type of **CN** ligand and one **NN** ligand using molSimplify v1.6.0 (ref. 55 and 56) and force field optimized again.

2.2. Feature sets

When developing machine learning models, it is important to strike the right balance between interpretability (*i.e.*, through features that relate to physical properties) and generalizability (*i.e.*, a model that performs well on complexes for which it was not trained). Thus, we evaluated and compared eight representations of the iridium complexes to identify the most suitable set of features for training ML models to predict Ir phosphor properties. The feature sets can be categorized into those based on substructure/fingerprints (*i.e.*, Morgan, Dice), those based on graph descriptors (*i.e.*, whole-complex revised autocorrelations, referred to as RACs,[57] ligand-only RACs, and Coulomb-decay RACs,[58] referred to as CD-RACs), and those based on electronic structure calculations (*i.e.*, xTB, ωPBEh, and B3LYP). For the substructure feature sets, we generated Morgan fingerprints,[59,60] which have been used previously in machine learning chemistry applications,[61–66] by one-hot encoding of groups of atoms in a structure. We computed these with a radius of three and 2048 bits on the isolated **CN** and **NN** ligands to capture the presence and absence of chemical substructures. We also generated Dice similarity coefficients[59] of ligand Morgan fingerprints. In this approach, we separately compare the Morgan fingerprints of the **CN** and **NN** ligand of each new iridium complex to all HLS **CN** or **NN** ligand Morgan fingerprints through the Dice similarity metric, which is a common measure to quantify the connectivity similarity of two molecules (ESI Text S2 and Table S3†). The Dice feature set size is determined by the number of training set HLS ligands (83 features in the random split, 78 in the grouped split, see next in Features for ANN models and performance). Dice similarity was selected after we found it outperforms the commonly employed Tanimoto similarity (ESI Table S4†).

Unlike similarity or fingerprint feature sets, graph-based representations capture the entire structure of the molecule, requiring any machine learning model trained on the graph representation to emphasize which components of the molecule matter most but with the potential benefit of generalizing to ligands that had not been seen before. For the graph-based feature sets, we generated RACs[57,67] for both the isolated ligands and the full iridium complex structures (ESI Text S2†). RACs are connectivity-based representations that have shown good performance for transition metal complex (TMC) property prediction.[43,57,68] For RACs, a TMC is represented as a molecular graph, with vertices for atoms and unweighted (*i.e.*, no bond length or order information) edges for bonds. Each RAC feature is the sum of products or the sum of differences of heuristic atomic properties at depth d on a TMC molecular graph, where d indicates the number of edges separating the starting and ending atoms (ESI Text S2†). The RACs include features that span the entire complex as well as weighted averages over the equatorial ligands and axial ligands, where **CN** and **NN** ligands may be classified as both when they are present in both the equatorial plane and axial position. We used the largest set of heuristic properties described in previous work, including both group number[69] and number of bonds,[58] leading to a final RAC feature set that contains 196 features (ESI Table S5†). For the ligand-only RAC feature set, we generated full-scope product RACs (*i.e.*, all atoms are used as starting atoms) on isolated **CN** and **NN** ligands for each TMC. We concatenated individual feature vectors for the **CN** ligands and **NN** ligands with equal weighting for each ligand type. The ligand-only RAC feature set contains significantly fewer (*i.e.*, only 70) total features than the RAC feature set (ESI Table S6†). We also generated Coulomb-decay RACs[70] on the iridium complex structures that were optimized with UFF (see Dataset). CD-RACs are a variant of RACs that also encode distances between the atoms in the RAC feature (ESI Text S2†). The CD-RAC feature set contains Coulomb-decay versions of the features in the RAC feature set but is of higher dimension (*i.e.*, 222 features) due to the added information from the geometry (ESI Table S7†).

Finally, we computed descriptors obtained from electronic structure theory, which were selected because they can be expected to correlate directly to the photophysical properties of the Ir complexes. Specifically, we selected electronic properties of the isolated ligands due to the lower computational cost in comparison to whole-complex properties. These ligand-based descriptors include the highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) energies of each ligand type, the ionization potential (IP) and electron affinity (EA) of each ligand type, and the partial charges (*i.e.*, Mulliken) of each of the metal-coordinating atoms (ESI Table S8†). For the xTB feature set, we utilized a specially reparametrized[71,72] vertical ionization potential and electron affinity-focused version of GFN1-xTB, a low-cost, semi-empirical tight binding method that has parameters for most elements in the periodic table.[73] We calculated ligand-only xTB features on UFF-optimized **CN** and **NN** ligands. The electronic structure features consist of quantum mechanical properties of the **CN** and **NN** ligands of a phosphor, and are in some cases correlated

to each other (*e.g.*, the HOMO and the EA of a ligand are closely related, ESI Table S8 and Fig. S2†). For the B3LYP and ωPBEh DFT feature sets, we performed density functional theory (DFT) calculations on isolated **CN** and **NN** ligands using the B3LYP[74–76] or ωPBEh[77] exchange correlation functionals respectively (see Computational details). Mulliken charges were used after they were found to outperform natural bond orbital (NBO) charges (ESI Table S9†).

3. Results and discussion

3.1. Features for ANN models and performance

The representation for the phosphor is a crucial piece in determining whether a machine learning (*i.e.*, ANN) model is likely to predict experimental properties accurately and to generalize to unseen complexes. A model and feature set that perform well on one property may perform poorly on another. We thus trained a total of 24 ANN models with each of the eight feature sets and the three target properties ($Em_{50/50}$, excited state lifetime, and emission spectral integral) and assessed their prediction performance on both a random split and grouped split of the training data. Here, random split refers to an 85/15 train/test partition, whereas in the grouped split five ligands are present only in the test set and are consequently unseen by the ANNs during training (see Computational details). The grouped split provides a more stringent test of how well our models generalize. For the random split train/test partition, the Dice, Morgan, and xTB feature sets lead to the lowest errors across all three target properties, suggesting they fit the data the best. The B3LYP DFT and RAC feature sets lead to the largest model test set errors, while the CD-RAC, ligand-only RAC, and ωPBEh DFT feature sets exhibit intermediate performance (Fig. 2 and ESI Fig. S3, S4 and Tables S10–S13†). The composition-based Dice and Morgan feature sets lead to the best predictions, judged on the basis of scaled MAEs of 0.03 to 0.05 for the three target properties (ESI Tables S11–S13†). There is a substantial difference in performance between feature sets: the percent difference in mean absolute error (MAE) on the test partition of the random split of the data between using the optimal feature set and the worst feature set for mean emission energy is 80%, and a similar performance erosion is observed for the other two properties (*i.e.*, 59% for spectral integral and 40% for phosphorescence lifetime), suggesting that it is important to select the best feature set to yield good performance.

We rationalize the relative performance of each feature set in the assessment on the random split of the training data by considering what aspects of ligands and complexes the different feature sets capture. We attribute the predictive power of the Dice feature set to the fact that phosphorescent properties are very ligand-directed, and in the random split of the training data, each ligand is represented in both the train and test sets. We attribute the predictive power of the Morgan feature set in the random split of the training data to the identification of substructures in ligands that affect phosphor excited state properties by tuning energy levels and ligand rigidity. Thus, features encoding ligand similarity to previously observed ligands and substructures present in ligands lead to the best performance on the random split. On the other end of the spectrum, the ANNs trained and assessed on a random split of the data using the RAC feature set (scaled MAE: 0.05 to 0.08) likely perform most poorly because they include significant metal-local information that does not vary across this set, because all complexes have an iridium center and an identical first coordination shell (ESI Tables S5 and S11–S13†). Thus, in datasets with a single metal center where only the ligands vary, standard similarity-based feature sets perform best in describing ligand variation. However, we still had to determine whether such feature sets are useful for discovery of novel complexes.

With regard to the ligand-only electronic structure feature sets, we surprisingly observe improved ANN performance on a random split of the training data with the xTB feature set

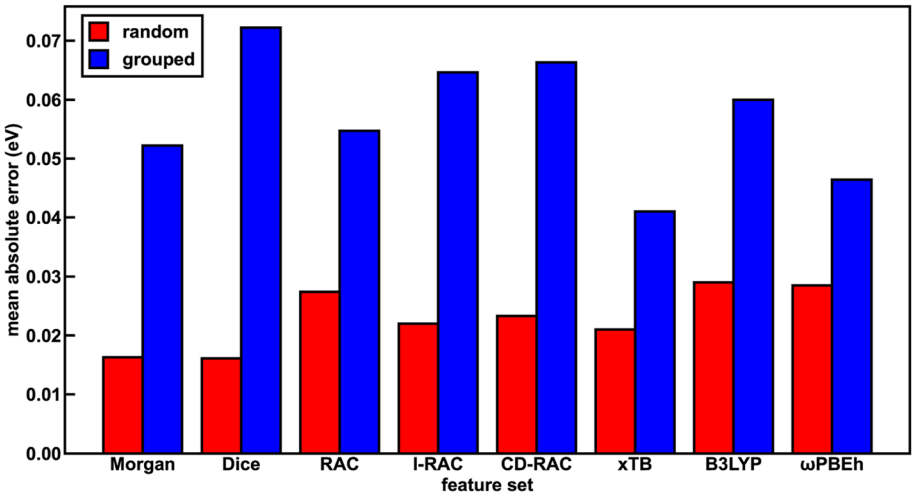

Fig. 2 The test set performance of ANNs trained on different feature sets in predicting $Em_{50/50}$ (MAE, in units of eV) for both random (red bars) and grouped splits (blue bars). Here, l-RAC refers to ligand-only RACs.

relative to the two DFT feature sets. While xTB is expected to be faster than DFT for feature generation, electronic structure properties from xTB alone should not necessarily be more accurate. The $Em_{50/50}$ xTB model error is 30% lower than that of the corresponding B3LYP DFT model (*i.e.*, 0.021 eV *vs.* 0.029 eV). Nevertheless, most (*i.e.*, eight of twelve) xTB features have high (>0.5) linear correlation with their B3LYP DFT and ωPBEh DFT counterparts, as determined by Pearson correlation coefficients (ESI Table S14†). Thus, the reparametrized GFN1-xTB method provides reliable electronic structure information from which our ANNs can generate accurate predictions. The reparametrized xTB method[71] is fitted to IP/EA values calculated with PW6B95/def2-TZVPD,[78] and it is possible that this functional and basis set combination achieve more accurate calculated properties than those generated in our DFT feature sets, leading to better ANN learning with the xTB feature set.

To assess the utility of ANN models trained on each of the feature sets for discovery of out-of-distribution complexes, we repeated the ANN training process on a grouped split, where five ligands are present only in the test set and are consequently unseen by the ANNs during training (see Computational details). We used the same grouped split across each property prediction task. We find that the test accuracy of the ANNs trained and tested on the grouped split of the data is worse than that of the corresponding ANNs trained and tested on a random split of the data for all features due to the presence of unseen ligands in the test set (ESI Tables S15–S17†). This worsened performance is most significant for the spectral integral and $Em_{50/50}$ target properties. Overall, the change in MAE averaged over all feature sets is significantly worse for these two properties (157% or 164% worse on average for spectral integral and $Em_{50/50}$, respectively) than for phosphorescence lifetime (28% worse, ESI Table S18†).

With the grouped split, the predictive power of the xTB feature set is improved relative to the other feature sets (Fig. 2 and ESI Fig. S3, S4†). For $Em_{50/50}$ prediction, the xTB feature set improves from the third-best feature set to the best feature set as a result of its scaled MAE increasing less than the best performers assessed on a random split of the training data (*i.e.*, 0.04 to 0.078 for xTB *versus* Dice 0.031 to 0.138, ESI Tables S18 and S19†). In practice, this means that the $Em_{50/50}$ xTB MAE doubles from 0.021 eV to 0.041 eV, while the Dice MAE nearly quadruples from 0.016 eV to 0.072 eV. We attribute this particularly worsened performance of the Dice feature set to the loss of information about variations in ligand chemistry because features describing similarity to held out ligands are no longer in the feature set for the grouped split (ESI Table S3†). The poor generalizability of the Dice feature set can also be attributed to the pseudo one-hot encoding of ligands *via* the similarity scores. The xTB features, in contrast, convey physical information that extrapolates beyond the ligands seen in the training data. We ultimately chose the xTB feature set for further analysis in evaluating hypothetical complexes because the xTB feature set has favorable performance on the grouped split for all three properties, indicating that the ANNs using the xTB feature set generalize well.

Beyond test set error, one challenge for applying ML models to novel complexes is the need to know how confident we should be in their predictions (*i.e.*, to quantify the uncertainty). To quantify ANN uncertainty in predictions for new phosphors outside of our initial training set, we use the latent space distance as a measure of how similar a new phosphor is to the complexes used to train the model.[79] The latent space is the last layer of an ANN, from which the final prediction is made *via* linear regression, and thus the distance in latent space of a new compound to training data should provide a representation of how different a new molecule is from training data according to the model. To confirm that this is a good measure of similarity that quantifies uncertainty for the current prediction task, we assessed the influence of latent space distance on test set prediction accuracy of the ANNs trained on a random split of the training data using xTB features as inputs. Following prior work,[79] we computed the average distance to ten nearest neighbors in the latent space formed by the training set and discarded predictions on any test set phosphor with an uncertainty quantification (UQ) metric exceeding the cutoff. A nearly monotonic decrease in average model error *versus* UQ cutoff suggests the possibility to control the error of predictions on new phosphors by discarding any prediction with a large UQ metric (Fig. 3 and ESI Fig. S5, S6†). Based on analysis of this UQ metric, we choose to avoid making model predictions on novel complexes when the distance in latent space is significantly larger than that typically observed on the random split test set (*i.e.*, more than two standard deviations above the mean, see New compound exploration). Starting from a rescaled UQ metric where the most distant test complex is assigned a value of 1.0, the cutoff is largest for the spectral integral ANN (*i.e.*,

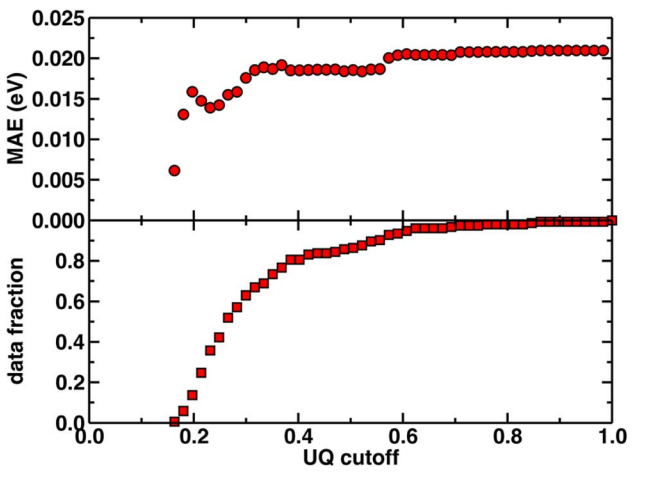

Fig. 3 The uncertainty quantification (UQ) cutoff *versus* test set mean absolute error (in eV) of the ANN model trained on a random split of the training data with the xTB feature set for predicting $Em_{50/50}$. The data fraction is the number of test set complexes under the corresponding UQ cutoff, and the MAE is calculated on this subset of complexes. The UQ metric used is the average latent space distance to the ten nearest neighbors in the training set following the protocol introduced in ref. 79. The UQ metric is normalized such that the largest UQ metric is scaled to 1.

0.79) and somewhat smaller for the lifetime and $Em_{50/50}$ ANNs (*i.e.*, 0.67 and 0.62).

3.2. Feature importance and trends

Given the high accuracy of the xTB-trained ANN models, we next sought to determine if simpler and more interpretable linear and random forest (*i.e.*, a series of binary decision trees) models trained on xTB features could attain similar accuracy. These models allow us to more transparently gain insight into which features most heavily influence phosphor property prediction. We trained random forest regression models that use xTB features to predict each of the three target properties. These random forest models have comparable performance to the ANNs and significantly outperform linear ridge regression models, (ESI Table S20 and Fig. S7†). Given the good performance of random forest models, we can analyze the most important features in these models to understand what features influence property prediction (*i.e.*, using impurity scores, Fig. 4). We find that xTB features of the **CN** ligand are more important than those of the **NN** ligand in predicting $Em_{50/50}$ and lifetime. For both of these target properties, the sum of impurity-based importances of **CN** ligand features is approximately 50% larger than the corresponding sum for **NN** ligand features, consistent with the presence of two **CN** ligands for each **NN** ligand in the complexes. The large role of the **CN** ligand in determining $Em_{50/50}$ can be explained by the partial localization of the phosphor complex HOMO on the **CN** ligand.[52] In contrast, xTB features of the **CN** and **NN** ligand are equally important in predicting the spectral integral. This indicates that when tuning $Em_{50/50}$ and lifetime, emphasis should be placed on selecting the **CN** ligand, whereas equal weight should be placed on varying **CN** and **NN** ligands to modify the spectral integral.

Focusing more on $Em_{50/50}$, we find that IP and EA are important for model predictions, as are the charges of metal-coordinating atoms (Fig. 4). Specifically, the top three xTB features for predicting $Em_{50/50}$ are the IP of the **CN** ligand and two of the coordinating nitrogen charges (*i.e.*, N charge (**CN**) and N1 charge (**NN**)). The importance of IP (**CN**) conforms to prior observations that ligand energy levels affect emission energy.[80,81] We also emphasize that these three xTB features vary significantly over the experimental dataset. The IP (**CN**) varies by nearly 1.5 eV (*i.e.*, from 7.56 eV to 9.03 eV), and the partial charges have a 0.1 a.u. range (*i.e.*, N charge (**CN**) from −0.35 a.u. to −0.24 a.u. and N1 charge (**NN**) from −0.37 a.u. to −0.28 a.u) (Fig. 5, 6 and ESI Fig. S8†). Thus, tuning these three features in a coordinated fashion should enable tuning of Ir phosphor complex $Em_{50/50}$.

As was observed from our global analysis, the most important xTB features are different for lifetime and spectral integral predictions (Fig. 4). For predicting lifetime, IP features from **CN** and **NN** ligands dominate, and the most important charge feature is the C charge of the **CN** ligand. For spectral integral, the top three features are EA (**NN**), IP (**NN**), and IP (**CN**), none of which are obtained from charges. The different feature importances for different target properties suggest some possibility of orthogonal design, wherein one phosphor property is tuned independently of the others. Nevertheless, given that ionization potential and electron affinity of the **CN** and **NN** ligands play

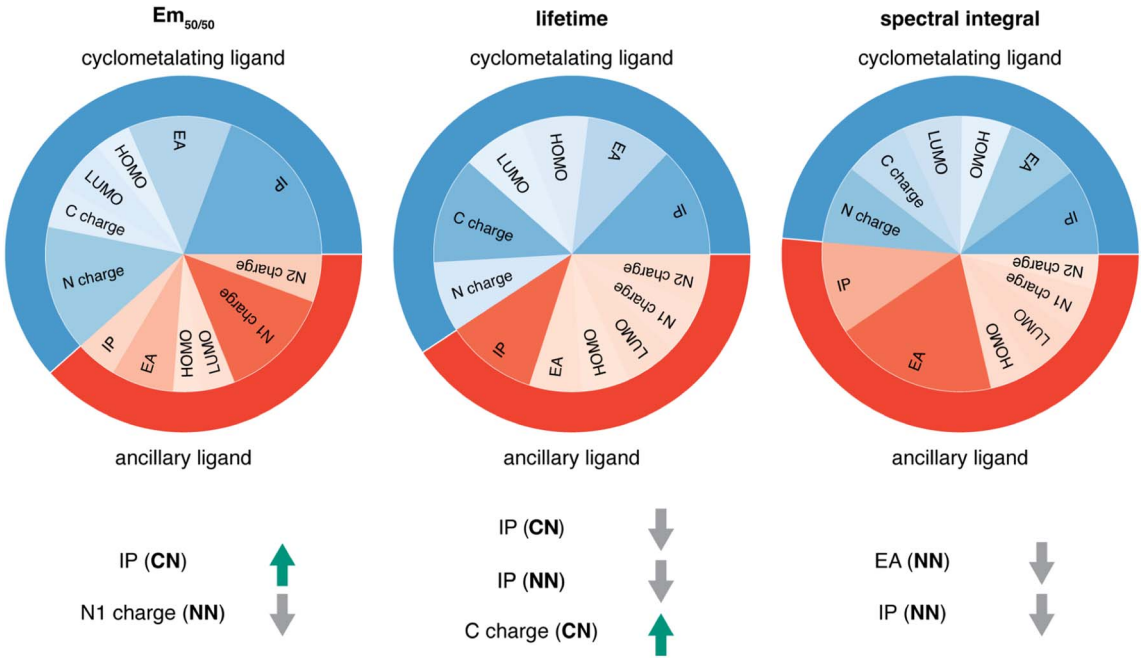

Fig. 4 For each of the three target properties, the corresponding column indicates: (top) random forest feature importances of the xTB **CN** and **NN** features and (bottom) the correlation of the most important xTB features to the target property, where a green arrow indicates positive correlation and a gray arrow indicates negative correlation. For example, IP (**CN**) is positively correlated to $Em_{50/50}$, while N1 charge (**NN**) is negatively correlated to $Em_{50/50}$.

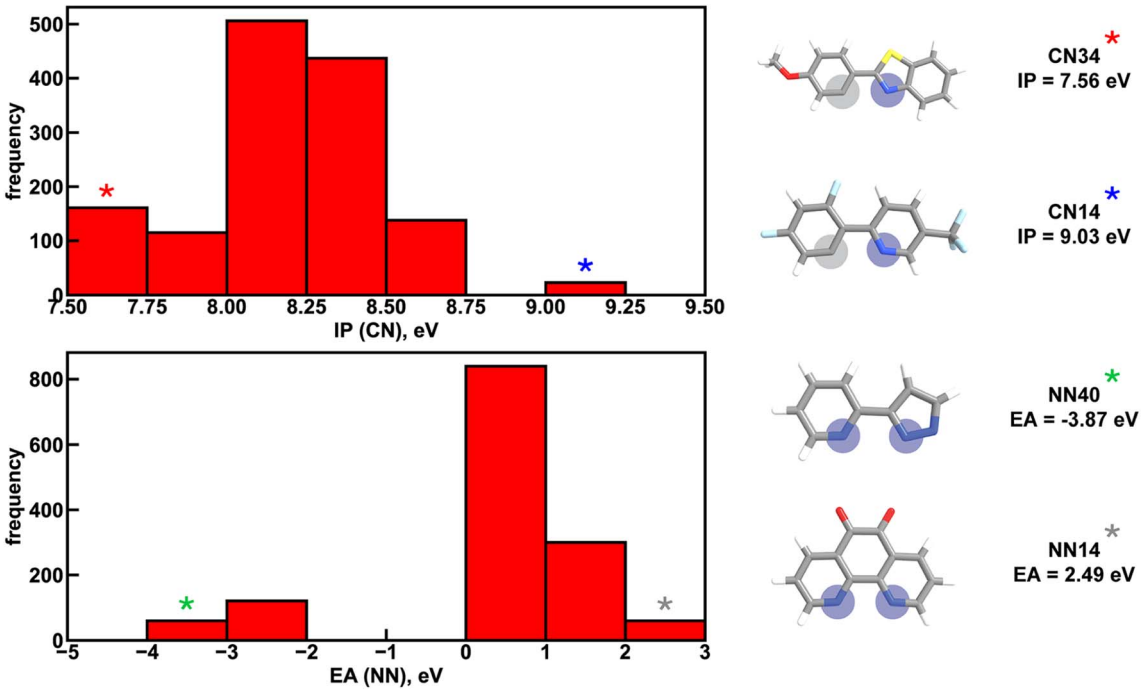

Fig. 5 Distribution of two xTB features across the experimental dataset of 1380 iridium phosphors. IP (**CN**) refers to the ionization potential of the **CN** ligand and EA (**NN**) refers to the electron affinity of the **NN** ligand. Asterisks correspond to ligands at the extreme ends of the distributions, shown on the right. Coordinating nitrogen (carbon) atoms are indicated with blue (gray) circles. Atoms are colored as follows: white for hydrogen, gray for carbon, blue for nitrogen, red for oxygen, light blue for fluorine, and yellow for sulfur.

a large role for all three target properties, altering coordinating atom charge without significantly altering the IP/EA is likely the most direct way to target changes in $Em_{50/50}$ or lifetime without altering the spectral integral.

Considering the most important xTB features as determined by random forest analysis, we further identified specific compounds with extreme (*i.e.*, high or low) experimental properties and compared how their xTB-computed features differed. For $Em_{50/50}$, high emission energy complexes typically have a high IP (**CN**), while low emission energy complexes typically have a low IP (**CN**). The N1 charge (**NN**) tends to be more positive for low emission energy complexes than for high emission energy ones. However, it is more challenging to identify which features are most important for long lifetime. In general, complexes with long lifetimes have a lower IP for both **CN** and **NN** ligands combined with a higher C charge (**CN**), but there are numerous exceptions. In the case of spectral integral, EA (**NN**) and IP (**NN**) are lower for bright complexes with high spectral integrals.

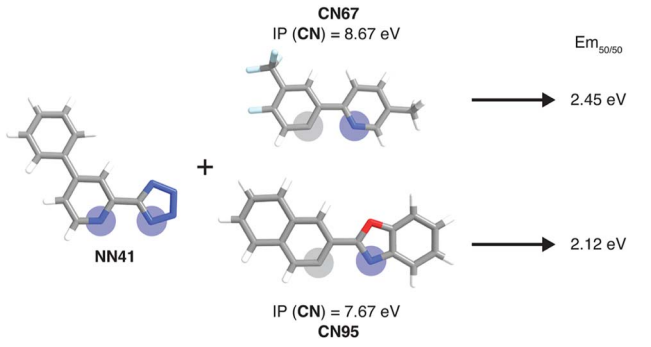

Fig. 6 Example of a pair of complexes where the substitution of the **CN** ligand leads to a large $Em_{50/50}$ property change. Coordinated nitrogen and carbon atoms are indicated with blue and gray circles, respectively. The relevant xTB features for the substituted ligands are shown. Atoms are colored as follows: white for hydrogen, gray for carbon, blue for nitrogen, red for oxygen, and light blue for fluorine.

To further identify specific examples of phosphors in the original experimental dataset that demonstrate the trends, we examined pairs of iridium complexes that differ only in the identity of one type of ligand. One such pair is $[Ir(\mathbf{CN67})_2(\mathbf{NN41})]^0$ and $[Ir(\mathbf{CN95})_2(\mathbf{NN41})]^0$ (Fig. 6). The former complex has an IP (**CN**) of 8.67 eV due to the electron-withdrawing fluorine groups on the cyclometalating ligand, while the latter complex has an IP (**CN**) of 7.67 eV. These values are on opposite ends of the IP (**CN**) distribution and contribute to $Em_{50/50}$ values on opposite ends of the $Em_{50/50}$ distribution, 2.45 eV and 2.12 eV, respectively (Fig. 5 and ESI Fig. S1†). The remaining five **CN** features for these two phosphors do not differ greatly from one another, underscoring the overriding effect of IP (**CN**). Similarly, increasing EA (**CN**) and EA (**NN**) can have a large effect on lifetime and spectral integral respectively (ESI Fig. S9 and S10†). These examples illustrate how differences in xTB features caused by ligand substitution correlate to shifts in phosphor properties.

To determine how ligand selection can allow for independent tuning, we consider the four complexes $[Ir(\mathbf{CN101})_2(\mathbf{NN2})]^+$, $[Ir(\mathbf{CN101})_2(\mathbf{NN20})]^+$, $[Ir(\mathbf{CN105})_2(\mathbf{NN2})]^+$, and $[Ir(\mathbf{CN105})_2(\mathbf{NN20})]^+$ that each differ by a single ligand. Changing the cyclometalating

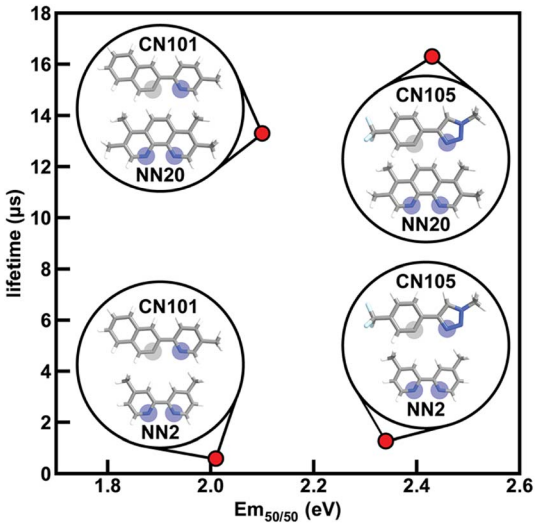

Fig. 7 Four iridium phosphor complexes and the effect of substituting the **CN** or **NN** ligand on $Em_{50/50}$ and lifetime indicated in the plot with structures shown as insets. Coordinated nitrogen (carbon) atoms are indicated with blue (gray) circles. Atoms are colored as follows: white for hydrogen, gray for carbon, blue for nitrogen, and light blue for fluorine.

ligand from **CN101** to **CN105** leads to an increase in $Em_{50/50}$ while having a small effect on phosphorescence lifetime, while changing the ancillary ligand from **NN2** to **NN20** leads to an increase in phosphorescence lifetime while having a small effect on $Em_{50/50}$ (Fig. 7). The increase in $Em_{50/50}$ when swapping **CN101** for **CN105** and the increase in lifetime when swapping **NN2** for **NN20** follows our observed trends of IP (**CN**) correlating positively to $Em_{50/50}$ and IP (**NN**) correlating negatively to lifetime. Furthermore, the small change in $Em_{50/50}$ when changing from **NN2** to **NN20** can be rationalized by the similar N1 charge (**NN**) between the two ancillary ligands. This example demonstrates how phosphor properties can be tuned orthogonally as guided by xTB features.

3.3. New compound exploration

We next aimed to demonstrate the utility of our ANNs in evaluating hypothetical complexes with ligands that were not in the training data but for which our models could make confident predictions. We applied one ANN for each property trained on a random split of the training data and used xTB features as inputs to screen hypothetical iridium complexes generated from CSD ligands (see Computational details and ESI Text S1†). Because the ANNs show better performance on random splits than grouped splits, they may be overfit to ligand chemistry present in the training data. Thus, we only considered hypothetical complexes under a UQ cutoff (*i.e.*, the distance in latent space) for all three ANNs. From a CSD screen, we identified 153 unique non-HLS **CN** ligands and 269 unique non-HLS **NN** ligands. Combining these new ligands with the HLS set led to 60 816 hypothetical complexes with at least one non-HLS ligand, of which 3598 hypothetical complexes fall within the UQ cutoff. This corresponds to inclusion of 70 unique non-HLS **CN** ligands and 42 unique non-HLS **NN** ligands in combination with each other or with HLS **CN** and **NN** ligands.

For this set of curated hypothetical complexes, we evaluated which ligands are present in the complexes with the highest and lowest ANN-predicted properties (ESI Fig. S11†). We find that specific ancillary ligands tend to be well-represented in complexes with extreme properties, indicating that phosphor properties are tuned by these ancillary ligands (ESI Table S21†). For example, the ligand that appears most often in hypothetical complexes with high predicted lifetime is the ancillary ligand from the CSD structure with refcode RASGAV. This conjugated ligand has a relatively low IP (**NN**) of 7.79 eV, which contributes to a longer lifetime following the previously identified trend (Fig. 4 and 8). Indeed, the other ancillary ligands that are well-represented in hypothetical complexes with extreme predicted lifetimes (**NN** ligands from complexes with refcodes FEQSEB, MIMYEO, TOTPAW, OVALEE, and MAXWIS) also follow the trend of low IP (**NN**) correlating to long lifetime (ESI Table S21†). We also note clear xTB feature trends in predictions for spectral integral and $Em_{50/50}$. The low IP (**NN**) of the ancillary ligand from RASGAV leads to a hypothetical complex with one of the highest predicted spectral integrals (Fig. 4, 8 and ESI Table S22†). With regard to $Em_{50/50}$, the fluorinated cyclometalating ligand from the CSD structure with refcode RADTEZ has a high ionization potential (9.24 eV). The high IP (**CN**) feature appears to contribute to a high emission energy, as the RADTEZ **CN** ligand is present in the three hypothetical complexes predicted to have the highest $Em_{50/50}$ values (Fig. 4, 8 and ESI Table S22†). On the other hand, the ancillary ligands LEZJAD **NN** and TUZHEE **NN** have high N1 charge (**NN**)

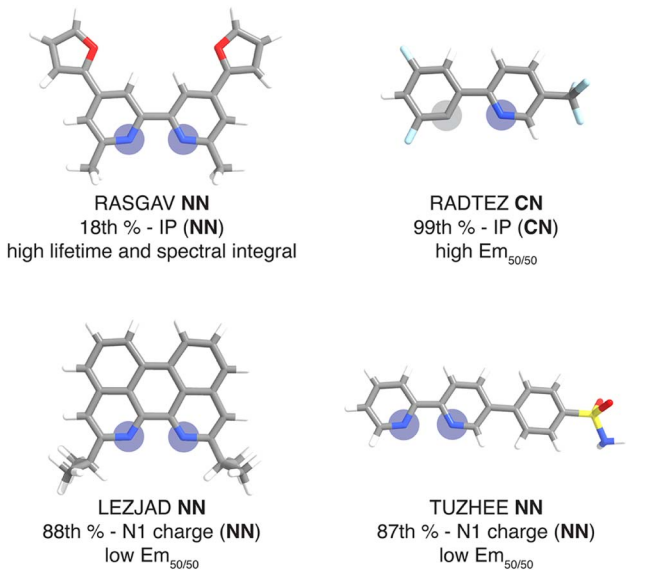

Fig. 8 Ligands mined from the CSD that lead to very high or very low phosphor properties predicted by the ANNs along with their percentile rank of the relevant property in the context of the experimental complexes. Coordinated nitrogen and carbon atoms are indicated with blue and gray circles respectively. Atoms are colored as follows: white for hydrogen, gray for carbon, blue for nitrogen, red for oxygen, light blue for fluorine, and yellow for sulfur.

features, leading to their presence in the three hypothetical complexes with the lowest predicted $Em_{50/50}$ values (Fig. 4, 8 and ESI Table S22†). Thus, we find that many ligands that lead to extreme hypothetical phosphor predicted properties follow our identified xTB feature trends from the experimental data. This lends interpretability to our model predictions and indicates that these predictions are derived from the electronic structure properties of the ligands.

To further validate performance of ANN models trained on a random split of the training data, we obtained TDDFT excited state energy and lifetime predictions and compared them to ANN predictions over complexes in both the experimental dataset and the uncertainty-controlled hypothetical dataset. Over a group of 26 representative test set complexes from the experimental dataset, we find that TDDFT overestimates the experimental emission energy by 0.3 eV on average, and further find that TDDFT predictions correlate with experiment less well than the $Em_{50/50}$ ANN predictions (Fig. 9 and ESI Tables S23–S25†). These results show that the $Em_{50/50}$ ANN achieves excellent performance. Even after applying a rigid downward shift to TDDFT energy predictions, they exhibit a larger spread around the experimental values than the predictions of our $Em_{50/50}$ ANN. Over the same 26 complexes, TDDFT lifetime predictions trend with experiment and ANN predictions; however, unlike the case of $Em_{50/50}$, TDDFT predictions outperform our lifetime ANN for complexes with long lifetimes (ESI Fig. S12†). This shortcoming of the lifetime ANN can be rationalized by the lower number of phosphors with long lifetimes in the experimental dataset used for model training (ESI Fig. S1 and Table S26†). Although we do not have ground-truth experimental data for our hypothetical set, we can still use TDDFT predictions to validate our ANN models. Over 21 representative hypothetical complexes, TDDFT energy predictions are on average 0.14 eV above $Em_{50/50}$ ANN predictions, and TDDFT lifetime predictions trend with lifetime ANN predictions (ESI Fig. S13–S15 and Tables S27, S28†). These results indicate that uncertainty-controlled ANN predictions over the hypothetical set of complexes are reliable, although they may underestimate the lifetime of phosphors with long lifetimes. Thus, as long as they are paired with suitable UQ metrics, the ANNs are trustworthy tools for the identification of hypothetical complexes with desired excited state properties.

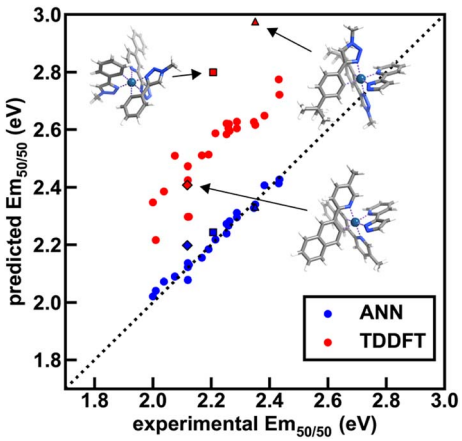

Fig. 9 Comparison of ANN and TDDFT $Em_{50/50}$ predictions to experiment (in eV) across 26 test set iridium complexes from the experimental dataset. These complexes were chosen to span the range of emission energies and lifetimes of the full set. TDDFT was carried out on optimized S_0 singlet geometries using the B3LYP functional, and the energies of the three lowest triplet sublevels were averaged to approximate $Em_{50/50}$; this approximation likely contributes to the worse performance of TDDFT relative to the ANN. Three high-error complexes ($[Ir(\mathbf{CN101})_2(\mathbf{NN40})]^0$, $[Ir(\mathbf{CN107})_2(\mathbf{NN41})]^0$, and $[Ir(\mathbf{CN109})_2(\mathbf{NN40})]^0$) are shown as insets, and their predicted and experimental $Em_{50/50}$ values are shown with black borders and unique shapes (diamond, square, and triangle, respectively). In the insets, atoms are colored as follows: white for hydrogen, gray for carbon, blue for nitrogen, and dark blue for iridium. The dotted line is included as a reference and corresponds to perfect agreement between prediction and experiment.

4. Conclusions

While *ab initio* methods like TDDFT are useful tools for studying excited states of iridium phosphors, they are computation-intensive and can also have insufficient accuracy, motivating the use of machine learning to leverage existing experimental data. Using experimental data on 1380 iridium phosphors, we trained ANNs to predict three experimental properties: $Em_{50/50}$, excited state lifetime, and emission spectral integral. We found that features calculated with xTB led to the best overall performance across the three properties on out-of-sample complexes, outperforming the standard Morgan fingerprint features and features based on RACs. We then used random forest regression models to determine which xTB features most influence phosphor properties and found that high cyclometalating ligand ionization potential is indicative of high $Em_{50/50}$, while high ancillary ligand ionization potential correlates to low lifetime and low spectral integral. These observations illustrate how phosphor properties can be altered through judicious ligand selection.

We next demonstrated how our ANNs can be applied to uncertainty-controlled chemical exploration by considering hypothetical iridium phosphors derived from ligands found in the CSD. We identified cyclometalating and ancillary ligands that lead to edge-of-distribution properties, such as an ancillary ligand predicted to result in both long-lifetime and high spectral integral phosphors. We confirmed the validity of these predictions by comparing to TDDFT, showing that for $Em_{50/50}$ the ANN significantly outperforms TDDFT, while for lifetime the corresponding ANN performs well only in regimes of sufficient training data. To improve the lifetime model predictions for long-lifetime complexes, further engineering of the features (*e.g.*, to incorporate non-local properties such as ligand flexibility) could improve performance. The ANN models for iridium phosphor property prediction that we present here are promising tools for chemical screening and the acceleration of chemical discovery, as they can be used to quickly evaluate thousands of hypothetical iridium phosphors to identify promising candidates for follow-up synthesis.

5. Computational details

5.1. Feature and structure generation

We generated feature sets to represent the 1380 Ir phosphor complexes as inputs to ML models (see Feature sets). We generated all ligands using the draw tool in Avogadro v1.1.2 (ref. 53 and 54) and subsequently optimized them with UFF.[82] We used molSimplify v1.6.0 for the generation of RAC feature sets on either ligands or complexes.[55,56] The xTB features were generated using xTB 6.4.0,[73] and DFT-based features were generated using the B3LYP[74–76] or ωPBEh[77] functional with the LACVP* basis set implemented in the TeraChem v1.9-2018.11-dev[83,84] program. For the generation of the Morgan and Dice feature sets, we used RDKit 2021.9.2 (ref. 85) both for Morgan fingerprints and the Dice similarity coefficients evaluated on Morgan fingerprints.

For DFT electronic structure descriptors, ligand geometries were geometry optimized with neutral charge and singlet spin multiplicity using DFT in TeraChem. Geometry optimizations used the L-BFGS algorithm in translation rotation internal coordinates (TRIC)[86] as implemented in TeraChem to the default tolerances of 4.5×10^{-4} hartree per bohr for the maximum gradient and 1×10^{-6} hartree for the change in energy between steps. We then performed single-point energy calculations on the optimized neutral ligand geometries at two different charges: +1 and −1 (ESI Text S1†). We used the LACVP* basis set, which for the HLS ligands corresponds to the LANL2DZ[87] effective core potential for Br and the 6-31G* basis set for all remaining elements. We calculated all non-singlet states with an unrestricted formalism and singlet states with a restricted formalism. Level shifting of 0.25 Ha was employed on both virtual and occupied orbitals to facilitate self-consistent field convergence. We used the hybrid DIIS[88]/A-DIIS[89] scheme for the self-consistent field procedure. We used TeraChem dynamic precision and a grid with approximately 3000 points per atom. Like the xTB feature set, DFT-generated features encode electronic structure information (ESI Table S8†).

5.2. ML models

We trained multiple artificial neural networks (ANNs) with each of the eight feature sets to predict three target properties: $Em_{50/50}$, excited state lifetime, and emission spectral integral. For all ANNs trained on a random split of the data, we used a random 70%/15%/15% train/validation/test split of the 1380 complexes from the prior study.[52] We find that results are robust to a random split with a larger test set allocation (56%/14%/30%, ESI Table S29†). To assess the generalizability of the ANNs, we also carried out grouped splits where we excluded from the training and validation data any complex containing a ligand from a select subset of **CN** and **NN** ligands. For the excluded ligands, we selected **CN21**, **CN103**, **CN104**, **NN20**, and **NN43** after determining these to be the most dissimilar HLS ligands relative to the other HLS ligands as measured through Dice similarities of Morgan fingerprints (ESI Tables S1, S2 and S30†). If one of the 1380 complexes contains one or more of these ligands, it is held out from the training set. We pre-processed features by normalizing each feature to a zero mean and unit variance over the train and validation data and removed any invariant features (ESI Text S2†). For ANNs predicting lifetime and $Em_{50/50}$, we excluded 356 complexes with low luminescent intensity (*i.e.* spectral integral less than 1×10^5 photon counts) from ANN training and performance evaluation due to the greater noise in lifetime and $Em_{50/50}$ measurements for dim Ir phosphors.

We built ANNs with Keras 2.4.3 with TensorFlow 2.3.0 as the backend.[90,91] Both bypass and residual layers were included as possible components of the ANN architecture for selection during hyperparameter optimization. Hyperparameters for each ANN were chosen using Hyperopt[92] with 200 evaluations, as judged by the mean absolute error of the model on the validation data. The built-in tree of Parzen estimator[93] algorithm in Hyperopt was used to select model hyperparameters. We used these chosen hyperparameters to train the final model on the combined train and validation data and evaluated performance on the test set (ESI Table S31†). All ANN models were trained with the AMSGrad variant[94] of the Adam optimizer[95] up to 2000 epochs. Dropout,[96] batch normalization,[97] and early stopping[98] were applied to avoid over-fitting. The patience for early stopping was 100. We enforced a floor of zero for all predictions since negative predictions for $Em_{50/50}$, lifetime, or spectral integral are unphysical. All machine learning models have been deposited online in a Zenodo repository.[99]

5.3. Out-of-distribution complexes

We identified hypothetical out-of-distribution iridium complexes which we enumerated combinatorially using **CN** and **NN** ligands not in the HLS. We selected these ligands by screening the CSD v5.42 + 2 updates, released in November 2020, for iridium complexes with two **CN** ligands and one **NN** ligand by specifying the first coordination sphere around iridium in a ConQuest 2021.1.0 search. Complexes selected by the screening were then examined by hand, and those that were not fit for analysis were eliminated (ESI Table S32†). The molSimplify code was used to identify unique ligands from the remaining complexes on the basis of their atom-weighted molecular graph determinants,[100] and we used any ligands not already in the HLS in combination with the HLS to generate new hypothetical $[Ir(\mathbf{CN})_2(\mathbf{NN})]^+$ complexes (ESI Text S1†).

5.4. TDDFT calculations

For *ab initio* validation of predictions using TDDFT, iridium phosphors were first geometry optimized, and TDDFT was then run on the optimized geometries using the ORCA 5.0.1 (ref. 101) program. All calculations employed a C-PCM solvation correction[102] to mimic DMSO. Singlet (*i.e.*, S_0) geometry optimization was carried out using the B3LYP[74–76] functional and the def2-TZVP[103] basis set with D4 dispersion correction[104] on structures generated by molSimplify. We found that using ground state singlet geometries instead of T_1 triplet geometries as inputs to TDDFT leads to better agreement with experiment, although geometries do not differ greatly in their RMSD (ESI Table S25 and Fig. S16†). Emission energies calculated with B3LYP were

found to correlate with experiment better than those calculated with the range-separated hybrid functionals CAM-B3LYP and ωB97X-D3BJ, motivating our use of B3LYP for TDDFT (ESI Fig. S17 and S18†). For TDDFT, the Zero-Order Regular Approximation (ZORA)[105] was used. The SARC-ZORA-TZVP[106] basis set was used for iridium and the ZORA-def2-TZVP basis set was used for all other elements along with the SARC/J auxiliary basis set. The TDDFT calculation included 25 roots. Quasi-degenerate perturbation theory spin–orbit coupling[107] was enabled, and the Tamm–Dancoff approximation was disabled.

Due to relativistic SOC caused by iridium, the T_1 manifold is split into three sublevels (zero-field splitting). For the *ab initio* calculation of emission energy, the energies of these three lowest triplet sublevels from the TDDFT calculation were averaged for each complex. For *ab initio* lifetime, radiative rate and radiative lifetime were calculated as in prior work[29–31,33,108] using output from TDDFT calculations. The radiative rate k_i from a triplet sublevel i is given by:

$$k_i = \frac{1}{\tau_i} = \frac{4}{3t_0}\alpha_0{}^3(\Delta E_i)^3 \sum_{\alpha\in\{x,y,z\}} \left|M_\alpha{}^i\right|^2 \quad (1)$$

where τ_i is the radiative lifetime of sublevel i, $t_0 = (4\pi\varepsilon_0)^2\hbar^3/m_e e^4$, α_0 is the fine structure constant, ΔE_i is the excitation energy in atomic units from the ground state to the sublevel i, and $M_\alpha{}^i$ is the α-axis projection of the transition dipole moment in atomic units between the ground state and the sublevel i.

The overall radiative lifetime from the three triplet sublevels is calculated as a Boltzmann average of radiative rates that depends on the energy differences between triplet sublevels.

$$\tau_{av} = \frac{1}{k_{av}} = \frac{1 + e^{-(\Delta E_{1,2}/k_B T)} + e^{-(\Delta E_{1,3}/k_B T)}}{k_1 + k_2 e^{-(\Delta E_{1,2}/k_B T)} + k_3 e^{-(\Delta E_{1,3}/k_B T)}} \quad (2)$$

$\Delta E_{1,2}$ is the energy difference between sublevels 1 and 2, and $\Delta E_{1,3}$ is the energy difference between sublevels 1 and 3. $T = 300$ K was used. This equation for lifetime does not take into account nonradiative decay, which can be significant in some cases. In order to account for the DMSO solvent, the calculated radiative lifetime was divided by the square of the refractive index of DMSO according to the Strickler–Berg relationship[109] in order to determine the final TDDFT lifetime.

Data availability

The datasets supporting this article have been uploaded as part of the ESI.† The ANN models associated with this work are deposited on Zenodo and have the following permanent DOI: **https://doi.org/10.5281/zenodo.7090416**.

Author contributions

Gianmarco G. Terrones: data curation, ML training, conceptualization, writing – original draft preparation, visualization; Chenru Duan: ML training, writing – reviewing and editing; Aditya Nandy: data curation, writing – reviewing and editing; Heather J. Kulik: writing – reviewing and editing, supervision, conceptualization.

Conflicts of interest

The authors declare no competing financial interest.

Acknowledgements

The authors acknowledge primary support for this work from the Office of Naval Research under grant numbers N00014-18-1-2434 and N00014-20-1-2150. Support for machine learning feature development was also provided by DARPA under grant number D18AP00039. G. G. T. was partially supported by an Alfred P. Sloan Foundation Scholarship (Grant Number G-2020-14067). A. N. was partially supported by the National Science Foundation Graduate Research Fellowship Program (Grant Number #1122374). This work was carried out in part using computational resources from the San Diego Supercomputer Cluster (SDSC), and in part using computational resources from the Extreme Science and Engineering Discovery Environment (XSEDE) which is supported by National Science Foundation grant number ACI-1548562. The authors acknowledge Adam H. Steeves for providing a critical reading of the manuscript.

References

1 V. C. Nikolis, A. Mischok, B. Siegmund, J. Kublitski, X. Jia, J. Benduhn, U. Hörmann, D. Neher, M. C. Gather, D. Spoltore, *et al.*, Strong Light-Matter Coupling for Reduced Photon Energy Losses in Organic Photovoltaics, *Nat. Commun.*, 2019, **10**, 1–8.
2 D. O. Hall and K. Rao, *Photosynthesis*, Cambridge University Press, 1999.
3 S. Tadepalli, J. M. Slocik, M. K. Gupta, R. R. Naik and S. Singamaneni, Bio-Optics and Bio-Inspired Optical Materials, *Chem. Rev.*, 2017, **117**, 12705–12763.
4 Y. Liang, D. Feng, Y. Wu, S.-T. Tsai, G. Li, C. Ray and L. Yu, Highly Efficient Solar Cell Polymers Developed via Fine-Tuning of Structural and Electronic Properties, *J. Am. Chem. Soc.*, 2009, **131**, 7792–7799.
5 W. Li, J. Fan, J. Li, Y. Mai and L. Wang, Controllable Grain Morphology of Perovskite Absorber Film by Molecular Self-Assembly toward Efficient Solar Cell Exceeding 17%, *J. Am. Chem. Soc.*, 2015, **137**, 10399–10405.
6 P. Li, Y. Zhou, Z. Zhao, Q. Xu, X. Wang, M. Xiao and Z. Zou, Hexahedron Prism-Anchored Octahedronal CeO_2: Crystal Facet-Based Homojunction Promoting Efficient Solar Fuel Synthesis, *J. Am. Chem. Soc.*, 2015, **137**, 9547–9550.
7 G. M. Farinola and R. Ragni, Electroluminescent Materials for White Organic Light Emitting Diodes, *Chem. Soc. Rev.*, 2011, **40**, 3467–3482.
8 G. Zhou, W.-Y. Wong and S. Suo, Recent Progress and Current Challenges in Phosphorescent White Organic Light-Emitting Diodes (WOLEDs), *J. Photochem. Photobiol., C*, 2010, **11**, 133–156.
9 T.-Y. Li, J. Wu, Z.-G. Wu, Y.-X. Zheng, J.-L. Zuo and Y. Pan, Rational Design of Phosphorescent Iridium(III) Complexes for Emission Color Tunability and Their Applications in OLEDs, *Coord. Chem. Rev.*, 2018, **374**, 55–92.

10 A. F. Henwood and E. Zysman-Colman, Luminescent Iridium Complexes Used in Light-Emitting Electrochemical Cells (LEECs), *Photoluminescent Materials and Electroluminescent Devices*, 2017, pp. 25–65.
11 C. K. Prier, D. A. Rankic and D. W. MacMillan, Visible Light Photoredox Catalysis with Transition Metal Complexes: Applications in Organic Synthesis, *Chem. Rev.*, 2013, **113**, 5322–5363.
12 V. Mdluli, S. Diluzio, J. Lewis, J. F. Kowalewski, T. U. Connell, D. Yaron, T. Kowalewski and S. Bernhard, High-Throughput Synthesis and Screening of Iridium(III) Photocatalysts for the Fast and Chemoselective Dehalogenation of Aryl Bromides, *ACS Catal.*, 2020, **10**, 6977–6987.
13 J. Lalevée, M. Peter, F. Dumur, D. Gigmes, N. Blanchard, M. A. Tehfe, F. Morlet-Savary and J. P. Fouassier, Subtle Ligand Effects in Oxidative Photocatalysis with Iridium Complexes: Application to Photopolymerization, *Chem.–Eur. J.*, 2011, **17**, 15027–15031.
14 S. Tobita and T. Yoshihara, Intracellular and in vivo Oxygen Sensing Using Phosphorescent Iridium(III) Complexes, *Curr. Opin. Chem. Biol.*, 2016, **33**, 39–45.
15 K. Y. Zhang, P. Gao, G. Sun, T. Zhang, X. Li, S. Liu, Q. Zhao, K. K.-W. Lo and W. Huang, Dual-Phosphorescent Iridium(III) Complexes Extending Oxygen Sensing from Hypoxia to Hyperoxia, *J. Am. Chem. Soc.*, 2018, **140**, 7827–7834.
16 H. Yersin, A. F. Rausch, R. Czerwieniec, T. Hofbeck and T. Fischer, The Triplet State of Organo-Transition Metal Compounds. Triplet Harvesting and Singlet Harvesting for Efficient OLEDs, *Coord. Chem. Rev.*, 2011, **255**, 2622–2652.
17 J.-H. Kim, S.-Y. Kim, Y.-J. Cho, H.-J. Son, D. W. Cho and S. O. Kang, A Detailed Evaluation for the Nonradiative Processes in Highly Phosphorescent Iridium(III) Complexes, *J. Phys. Chem. C*, 2018, **122**, 4029–4036.
18 C. You, D. Liu, J. Yu, H. Tan, M. Zhu, B. Zhang, Y. Liu, Y. Wang and W. Zhu, Boosting Efficiency of Near-Infrared Emitting Iridium (III) Phosphors by Administrating Their π–π Conjugation Effect of Core–Shell Structure in Solution-Processed OLEDs, *Adv. Opt. Mater.*, 2020, **8**, 2000154.
19 C. H. Yang, Y. M. Cheng, Y. Chi, C. J. Hsu, F. C. Fang, K. T. Wong, P. T. Chou, C. H. Chang, M. H. Tsai and C. C. Wu, Blue-Emitting Heteroleptic Iridium (III) Complexes Suitable for High-Efficiency Phosphorescent OLEDs, *Angew. Chem.*, 2007, **119**, 2470–2473.
20 Y. Liu, G. Gahungu, X. Sun, J. Su, X. Qu and Z. Wu, Theoretical Study on the Influence of Ancillary and Cyclometalated Ligands on the Electronic Structures and Optoelectronic Properties of Heteroleptic Iridium(III) Complexes, *Dalton Trans.*, 2012, **41**, 7595–7603.
21 K. Świderek and P. Paneth, Modeling Excitation Properties of Iridium Complexes, *J. Phys. Org. Chem.*, 2009, **22**, 845–856.
22 X. Li, Q. Zhang, Y. Tu, H. Ågren and H. Tian, Modulation of Iridium(III) Phosphorescence via Photochromic Ligands: A Density Functional Theory Study, *Phys. Chem. Chem. Phys.*, 2010, **12**, 13730–13736.
23 X. Li, B. Minaev, H. Ågren and H. Tian, Theoretical Study of Phosphorescence of Iridium Complexes with Fluorine-Substituted Phenylpyridine Ligands, *Eur. J. Inorg. Chem.*, 2011, 2517–2524.
24 Y. Liu, G. Gahungu, X. Sun, X. Qu and Z. Wu, Effects of N-Substitution on Phosphorescence Efficiency and Color Tuning of a Series of Ir(III) Complexes with a Phosphite Tripod Ligand: A DFT/TDDFT Study, *J. Phys. Chem. C*, 2012, **116**, 26496–26506.
25 J. M. Younker and K. D. Dobbs, Correlating Experimental Photophysical Properties of Iridium(III) Complexes to Spin–Orbit Coupled TDDFT Predictions, *J. Phys. Chem. C*, 2013, **117**, 25714–25723.
26 F. Monti, A. Baschieri, I. Gualandi, J. J. Serrano-Pérez, J. M. Junquera-Hernández, D. Tonelli, A. Mazzanti, S. Muzzioli, S. Stagni, C. Roldan-Carmona, *et al.*, Iridium(III) Complexes with Phenyl-Tetrazoles as Cyclometalating Ligands, *Inorg. Chem.*, 2014, **53**, 7709–7721.
27 K. P. Zanoni, B. K. Kariyazaki, A. Ito, M. K. Brennaman, T. J. Meyer and N. Y. Murakami Iha, Blue-Green Iridium(III) Emitter and Comprehensive Photophysical Elucidation of Heteroleptic Cyclometalated Iridium(III) Complexes, *Inorg. Chem.*, 2014, **53**, 4089–4099.
28 S. Fantacci and F. De Angelis, A Computational Approach to the Electronic and Optical Properties of Ru (II) and Ir (III) Polypyridyl Complexes: Applications to DSC, OLED and NLO, *Coord. Chem. Rev.*, 2011, **255**, 2704–2726.
29 E. Jansson, B. Minaev, S. Schrader and H. Ågren, Time-dependent density functional calculations of phosphorescence parameters for fac-tris (2-phenylpyridine) iridium, *Chem. Phys.*, 2007, **333**, 157–167.
30 B. Minaev, V. Minaeva and H. Ågren, Theoretical Study of the Cyclometalated Iridium(III) Complexes Used as Chromophores for Organic Light-Emitting Diodes, *J. Phys. Chem. A*, 2009, **113**, 726–735.
31 A. R. Smith, P. L. Burn and B. J. Powell, Spin–Orbit Coupling in Phosphorescent Iridium(III) Complexes, *ChemPhysChem*, 2011, **12**, 2429–2438.
32 A. Smith, M. Riley, P. Burn, I. Gentle, S.-C. Lo and B. Powell, Effects of Fluorination on Iridium(III) Complex Phosphorescence: Magnetic Circular Dichroism and Relativistic Time-Dependent Density Functional Theory, *Inorg. Chem.*, 2012, **51**, 2821–2831.
33 K. Mori, T. Goumans, E. Van Lenthe and F. Wang, Predicting Phosphorescent Lifetimes and Zero-Field Splitting of Organometallic Complexes with Time-Dependent Density Functional Theory Including Spin–Orbit Coupling, *Phys. Chem. Chem. Phys.*, 2014, **16**, 14523–14530.
34 Q. Peng, Y. Niu, Q. Shi, X. Gao and Z. Shuai, Correlation Function Formalism for Triplet Excited State Decay: Combined Spin–Orbit and Nonadiabatic Couplings, *J. Chem. Theory Comput.*, 2013, **9**, 1132–1143.

35 Q. Peng, Q. Shi, Y. Niu, Y. Yi, S. Sun, W. Li and Z. Shuai, Understanding the Efficiency Drooping of the Deep Blue Organometallic Phosphors: A Computational Study of Radiative and Non-Radiative Decay Rates for Triplets, *J. Mater. Chem. C*, 2016, **4**, 6829–6838.
36 X. Zhang, D. Jacquemin, Q. Peng, Z. Shuai and D. Escudero, General Approach to Compute Phosphorescent OLED Efficiency, *J. Phys. Chem. C*, 2018, **122**, 6340–6347.
37 D. Escudero, Quantitative Prediction of Photoluminescence Quantum Yields of Phosphors from First Principles, *Chem. Sci.*, 2016, **7**, 1262–1267.
38 B. Mortazavi, E. V. Podryabinkin, I. S. Novikov, T. Rabczuk, X. Zhuang and A. V. Shapeev, Accelerating First-Principles Estimation of Thermal Conductivity by Machine-Learning Interatomic Potentials: A MTP/ShengBTE Solution, *Comput. Phys. Commun.*, 2021, **258**, 107583.
39 J. Behler, Neural Network Potential-Energy Surfaces in Chemistry: A Tool for Large-Scale Simulations, *Phys. Chem. Chem. Phys.*, 2011, **13**, 17930–17955.
40 V. Botu and R. Ramprasad, Adaptive Machine Learning Framework to Accelerate ab initio Molecular Dynamics, *Int. J. Quantum Chem.*, 2015, **115**, 1074–1083.
41 J. Westermayr and P. Marquetand, Machine Learning for Electronically Excited States of Molecules, *Chem. Rev.*, 2020, **121**, 9873–9926.
42 N. Fey and J. M. Lynam, Computational Mechanistic Study in Organometallic Catalysis: Why Prediction Is Still a Challenge, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2022, **12**, e1590.
43 J. P. Janet, S. Ramesh, C. Duan and H. J. Kulik, 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.
44 A. Nandy, C. Duan and H. J. Kulik, Using Machine Learning and Data Mining to Leverage Community Knowledge for the Engineering of Stable Metal–Organic Frameworks, *J. Am. Chem. Soc.*, 2021, **143**, 17535–17547.
45 A. E. Sifain, L. Lystrom, R. A. Messerly, J. S. Smith, B. Nebgen, K. Barros, S. Tretiak, N. Lubbers and B. J. Gifford, Predicting Phosphorescence Energies and Inferring Wavefunction Localization with Machine Learning, *Chem. Sci.*, 2021, **12**, 10207–10217.
46 R. Gómez-Bombarelli, J. Aguilera-Iparraguirre, T. D. Hirzel, D. Duvenaud, D. Maclaurin, M. A. Blood-Forsythe, H. S. Chae, M. Einzinger, D.-G. Ha, T. Wu, *et al.*, Design of Efficient Molecular Organic Light-Emitting Diodes by a High-Throughput Virtual Screening and Experimental Approach, *Nat. Mater.*, 2016, **15**, 1120–1127.
47 C.-W. Ju, H. Bai, B. Li and R. Liu, Machine Learning Enables Highly Accurate Predictions of Photophysical Properties of Organic Fluorescent Materials: Emission Wavelengths and Quantum Yields, *J. Chem. Inf. Model.*, 2021, **61**, 1053–1065.
48 P. Friederich, G. dos Passos Gomes, R. De Bin, A. Aspuru-Guzik and D. Balcells, Machine Learning Dihydrogen Activation in the Chemical Space Surrounding Vaska's Complex, *Chem. Sci.*, 2020, **11**, 4584–4601.
49 H. Chen, S. Yamaguchi, Y. Morita, H. Nakao, X. Zhai, Y. Shimizu, H. Mitsunuma and M. Kanai, Data-Driven Catalyst Optimization for Stereodivergent Asymmetric Synthesis by Iridium/Boron Hybrid Catalysis, *Cell Rep. Phys. Sci.*, 2021, **2**, 100679.
50 G. Terrones, C. Duan, A. Nandy and H. J. Kulik, Low-cost machine learning approach to the prediction of transition metal phosphor excited state properties, *arXiv*, 2022, preprint, arXiv:2209.08595.
51 A. Karuth, G. Casanola-Martin, L. Lystrom, W. Sun, D. Kilin, S. Kilina and B. Rasulev, *Combined Machine Learning, Computational and Experimental Analysis of the Iridium (III) Complexes with Red to Near-IR Emission*, 2022.
52 S. DiLuzio, V. Mdluli, T. U. Connell, J. Lewis, V. VanBenschoten and S. Bernhard, High-Throughput Screening and Automated Data-Driven Analysis of the Triplet Photophysical Properties of Structurally Diverse, Heteroleptic Iridium(III) Complexes, *J. Am. Chem. Soc.*, 2021, **143**, 1179–1194.
53 *Avogadro: An Open-Source Molecular Builder and Visualization Tool. Version 1.1.2*, **https://avogadro.cc/**.
54 M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek and G. R. Hutchison, Avogadro: An Advanced Semantic Chemical Editor, Visualization, and Analysis Platform, *J. Cheminf.*, 2012, **4**, 1–17.
55 E. I. Ioannidis, T. Z. Gani and H. J. Kulik, molSimplify: A Toolkit for Automating Discovery in Inorganic Chemistry, *J. Comput. Chem.*, 2016, **37**, 2106–2117.
56 J. P. Janet, T. Z. Gani, A. H. Steeves, E. I. Ioannidis and H. J. Kulik, Leveraging Cheminformatics Strategies for Inorganic Discovery: Application to Redox Potential Design, *Ind. Eng. Chem. Res.*, 2017, **56**, 4898–4910.
57 J. P. Janet and H. J. Kulik, Resolving Transition Metal Chemical Space: Feature Selection for Machine Learning and Structure–Property Relationships, *J. Phys. Chem. A*, 2017, **121**, 8939–8954.
58 C. Duan, F. Liu, A. Nandy and H. J. Kulik, Data-Driven Approaches Can Overcome the Cost-Accuracy Trade-Off in Multireference Diagnostics, *J. Chem. Theory Comput.*, 2020, **16**, 4373–4387.
59 A. Cereto-Massagué, M. J. Ojeda, C. Valls, M. Mulero, S. Garcia-Vallvé and G. Pujadas, Molecular Fingerprint Similarity Search in Virtual Screening, *Methods*, 2015, **71**, 58–63.
60 D. Rogers and M. Hahn, Extended-Connectivity Fingerprints, *J. Chem. Inf. Model.*, 2010, **50**, 742–754.
61 B. C. Barnes, D. C. Elton, Z. Boukouvalas, D. E. Taylor, W. D. Mattson, M. D. Fuge and P. W. Chung, Machine Learning of Energetic Material Properties, *arXiv*, 2018, preprint, arXiv:1807.06156, DOI: **10.48550/arXiv.1807.06156**.
62 D. C. Elton, Z. Boukouvalas, M. S. Butrico, M. D. Fuge and P. W. Chung, Applying Machine Learning Techniques to Predict the Properties of Energetic Materials, *Sci. Rep.*, 2018, **8**, 1–12.
63 L. Tao, G. Chen and Y. Li, Machine Learning Discovery of High-Temperature Polymers, *Patterns*, 2021, **2**, 100225.

64 F. O. Sanches-Neto, J. R. Dias-Silva, L. H. Keng Queiroz Junior and V. H. Carvalho-Silva, "py SiRC": Machine Learning Combined with Molecular Fingerprints to Predict the Reaction Rate Constant of the Radical-Based Oxidation Processes of Aqueous Organic Contaminants, *Environ. Sci. Technol.*, 2021, **55**, 12437–12448.

65 K. Jorner, T. Brinck, P.-O. Norrby and D. Buttar, Machine Learning Meets Mechanistic Modelling for Accurate Prediction of Experimental Activation Energies, *Chem. Sci.*, 2021, **12**, 1163–1175.

66 T. Fujimoto and H. Gotoh, Prediction and Chemical Interpretation of Singlet-Oxygen-Scavenging Activity of Small Molecule Compounds by Using Machine Learning, *Antioxidants*, 2021, **10**, 1751.

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

68 A. Nandy, J. Zhu, J. P. Janet, C. Duan, R. B. Getman and H. J. Kulik, Machine Learning Accelerates the Discovery of Design Rules and Exceptions in Stable Metal–Oxo Intermediate Formation, *ACS Catal.*, 2019, **9**, 8243–8255.

69 D. R. Harper, A. Nandy, N. Arunachalam, C. Duan, J. P. Janet and H. J. Kulik, Representations and Strategies for Transferable Machine Learning Improve Model Performance in Chemical Discovery, *J. Chem. Phys.*, 2022, **156**, 074101.

70 C. Duan, F. Liu, A. Nandy and H. J. Kulik, Data-Driven Approaches Can Overcome the Cost–Accuracy Trade-Off in Multireference Diagnostics, *J. Chem. Theory Comput.*, 2020, **16**, 4373–4387.

71 V. Ásgeirsson, C. A. Bauer and S. Grimme, Quantum Chemical Calculation of Electron Ionization Mass Spectra for General Organic and Inorganic Molecules, *Chem. Sci.*, 2017, **8**, 4879–4895.

72 S. Grimme, *Vertical Ionization Potentials and Electron Affinities*, **https://xtb-docs.readthedocs.io/en/latest/sp.html#vertical-ionization-potentials-and-electron-affinities**, accessed October 20, 2022.

73 S. Grimme, C. Bannwarth and P. Shushkov, A Robust and Accurate Tight-Binding Quantum Chemical Method for Structures, Vibrational Frequencies, and Noncovalent Interactions of Large Molecular Systems Parametrized for All spd-Block Elements (Z= 1–86), *J. Chem. Theory Comput.*, 2017, **13**, 1989–2009.

74 P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, Ab initio Calculation of Vibrational Absorption and Circular Dichroism Spectra using Density Functional Force Fields, *J. Phys. Chem.*, 1994, **98**, 11623–11627.

75 A. D. Beck, Density-Functional Thermochemistry. III. The Role of Exact Exchange, *J. Chem. Phys.*, 1993, **98**, 5648.

76 C. Lee, W. Yang and R. G. Parr, Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1988, **37**, 785.

77 J. Heyd, G. E. Scuseria and M. Ernzerhof, Hybrid Functionals Based on a Screened Coulomb Potential, *J. Chem. Phys.*, 2003, **118**, 8207–8215.

78 Y. Zhao and D. G. Truhlar, Design of Density Functionals that are Broadly Accurate for Thermochemistry, Thermochemical Kinetics, and Nonbonded Interactions, *J. Phys. Chem. A*, 2005, **109**, 5656–5667.

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

80 J.-H. Kim, S.-Y. Kim, S. Jang, S. Yi, D. W. Cho, H.-J. Son and S. O. Kang, Blue Phosphorescence with High Quantum Efficiency Engaging the Trifluoromethylsulfonyl Group to Iridium Phenylpyridine Complexes, *Inorg. Chem.*, 2019, **58**, 16112–16125.

81 K. Hasan, A. K. Bansal, I. D. Samuel, C. Roldán-Carmona, H. J. Bolink and E. Zysman-Colman, Tuning the Emission of Cationic Iridium(III) Complexes Towards the Red through Methoxy Substitution of the Cyclometalating Ligand, *Sci. Rep.*, 2015, **5**, 1–16.

82 A. K. Rappé, C. J. Casewit, K. Colwell, W. A. Goddard III and W. M. Skiff, UFF, a Full Periodic Table Force Field for Molecular Mechanics and Molecular Dynamics Simulations, *J. Am. Chem. Soc.*, 1992, **114**, 10024–10035.

83 *Petachem*, **https://www.petachem.com/**, accessed October 20, 2022.

84 S. Seritan, C. Bannwarth, B. S. Fales, E. G. Hohenstein, S. I. Kokkila-Schumacher, N. Luehr, J. W. Snyder Jr, C. Song, A. V. Titov, I. S. Ufimtsev, *et al.*, TeraChem: Accelerating Electronic Structure and ab initio Molecular Dynamics with Graphical Processing Units, *J. Chem. Phys.*, 2020, **152**, 224110.

85 G. Landrum, RDKit Documentation, *Release*, 2013, **1**, 4.

86 L.-P. Wang and C. Song, Geometry Optimization Made Simple with Translation and Rotation Coordinates, *J. Chem. Phys.*, 2016, **144**, 214108.

87 P. J. Hay and W. R. Wadt, Ab initio Effective Core Potentials for Molecular Calculations. Potentials for the Transition Metal Atoms Sc to Hg, *J. Chem. Phys.*, 1985, **82**, 270–283.

88 P. Pulay, Improved SCF Convergence Acceleration, *J. Comput. Chem.*, 1982, **3**, 556–560.

89 X. Hu and W. Yang, Accelerating Self-Consistent Field Convergence with the Augmented Roothaan–Hall Energy Function, *J. Chem. Phys.*, 2010, **132**, 054109.

90 *TensorFlow*, **https://www.tensorflow.org/**, accessed October 20, 2022.

91 M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, *et al.*, Tensorflow: Large-scale machine learning on heterogeneous distributed systems, *arXiv*, 2016, preprint, arXiv:1603.04467, DOI: **10.48550/arXiv.1603.04467**.

92 J. Bergstra, B. Komer, C. Eliasmith, D. Yamins and D. D. Cox, Hyperopt: A Python Library for Model Selection and Hyperparameter Optimization, *Comput. Sci. Discovery*, 2015, **8**, 014008.

93 J. Bergstra, R. Bardenet, Y. Bengio and B. Kégl, Algorithms for Hyper-Parameter Optimization, *Adv. Neural Inf. Process. Syst.*, 2011, **24**, 2546–2554.
94 S. J. Reddi, S. Kale and S. Kumar, On the Convergence of Adam and Beyond, *arXiv*, 2019, preprint, arXiv:1904.09237, DOI: **10.48550/arXiv.1904.09237**.
95 D. P. Kingma and J. Ba, Adam: A Method for Stochastic Optimization, *arXiv*, 2014, preprint, arXiv:1412.6980, DOI: **10.48550/arXiv.1412.6980**.
96 N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever and R. Salakhutdinov, Dropout: A Simple Way to Prevent Neural Networks from Overfitting, *J. Mach. Learn. Res.*, 2014, **15**, 1929–1958.
97 S. Ioffe and C. Szegedy, *International conference on machine learning*, 2015, pp. 448–456.
98 R. Caruana, S. Lawrence and L. Giles, Overfitting in Neural Nets: Backpropagation, Conjugate Gradient, and Early Stopping, *Adv. Neural Inf. Process. Syst.*, 2000, **13**, 402–408.
99 *Zenodo Repository for Low-Cost Machine Learning Prediction of Excited State Properties of Iridium-Centered Phosphors*, **https://zenodo.org/record/7090417**, accessed October 20, 2022.
100 M. G. Taylor, T. Yang, S. Lin, A. Nandy, J. P. Janet, C. Duan and H. J. Kulik, Seeing is Believing: Experimental Spin States from Machine Learning Model Structure Predictions, *J. Phys. Chem. A*, 2020, **124**, 3286–3299.
101 F. Neese, The ORCA Program System, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2012, **2**, 73–78.
102 V. Barone and M. Cossi, Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model, *J. Phys. Chem. A*, 1998, **102**, 1995–2001.
103 F. Weigend and R. Ahlrichs, Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3297–3305.
104 E. Caldeweyher, S. Ehlert, A. Hansen, H. Neugebauer, S. Spicher, C. Bannwarth and S. Grimme, A Generally Applicable Atomic-Charge Dependent London Dispersion Correction, *J. Chem. Phys.*, 2019, **150**, 154122.
105 C. van Wüllen, Molecular Density Functional Calculations in the Regular Relativistic Approximation: Method, Application to Coinage Metal Diatomics, Hydrides, Fluorides and Chlorides, and Comparison with First-Order Relativistic Calculations, *J. Chem. Phys.*, 1998, **109**, 392–399.
106 D. A. Pantazis, X.-Y. Chen, C. R. Landis and F. Neese, All-Electron Scalar Relativistic Basis Sets for Third-Row Transition Metal Atoms, *J. Chem. Theory Comput.*, 2008, **4**, 908–919.
107 B. de Souza, G. Farias, F. Neese and R. Izsák, Predicting Phosphorescence Rates of Light Organic Molecules Using Time-Dependent Density Functional Theory and the Path Integral Approach to Dynamics, *J. Chem. Theory Comput.*, 2019, **15**, 1896–1904.
108 I. Soriano-Díaz, E. Ortí and A. Giussani, On the Importance of Ligand-Centered Excited States in the Emission of Cyclometalated Ir (III) Complexes, *Inorg. Chem.*, 2021, **60**, 13222–13232.
109 S. Strickler and R. A. Berg, Relationship between Absorption Intensity and Fluorescence Lifetime of Molecules, *J. Chem. Phys.*, 1962, **37**, 814–822.