

Predicting Emission Spectra of Heteroleptic Iridium Complexes Using Artificial Chemical Intelligence

Yudhajit Pal,[†] Tahoe A. Fiala,[‡] Wesley B. Swords,[‡] Tehshik P. Yoon,[‡] and J.R. Schmidt^{*,†}

[†]*Theoretical Chemistry Institute and Department of Chemistry, 1101 University Avenue,
University of Wisconsin-Madison, Madison, WI 53706, United States*

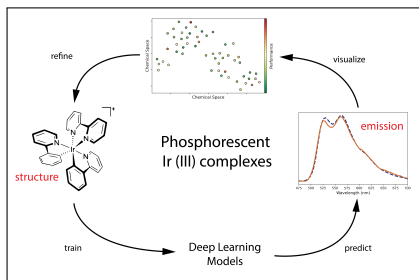
[‡]*Department of Chemistry, 1101 University Avenue, University of Wisconsin-Madison,
Madison, WI 53706, United States*

E-mail: schmidt@chem.wisc.edu

Abstract

We report a deep learning-based approach to accurately predict the emission spectra of phosphorescent heteroleptic $[\text{Ir}(\text{C}^{\wedge}\text{N})_2(\text{N}^{\wedge}\text{N})]^+$ complexes, enabling the rapid discovery of novel Ir(III) chromophores for diverse applications including organic light-emitting diodes and solar fuel cells. The deep learning models utilize graph neural networks and other chemical features in architectures that reflect the inherent structure of the heteroleptic complexes, composed of $\text{C}^{\wedge}\text{N}$ and $\text{N}^{\wedge}\text{N}$ ligands, and are thus geared towards efficient training over the dataset. By leveraging experimental emission data, our models reliably predict the full emission spectra of these complexes across various emission profiles, surpassing the accuracy of conventional DFT and correlated wavefunction methods, while simultaneously achieving robustness to the presence of imperfect (noisy, low-quality) training spectra. We showcase the potential applications for these and related models for *in silico* prediction of complexes with tailored emission properties, as well as in “design of experiment” contexts to reduce the synthetic burden of high-throughput screening. In the latter case, we demonstrate that the models allow us to exploit a limited amount of experimental data to explore a wide range of chemical space, thus leveraging a modest synthetic effort.

TOC Graphic



Introduction

Transition metal complexes of d^6 second- and third-row transition metal elements can exhibit phosphorescence properties ideal for a range of important applications. In particular, cyclometalated octahedral Ir(III) complexes are common elements of organic light-emitting diodes¹⁻⁵ and solar fuel cells.⁶ They have also increasingly become utilized in organic photocatalytic reactions,^{1,7-9} photodynamic therapy,^{10,11} and bioimaging applications.¹² The utility of Ir(III) chromophores across so many different applications arises from several attractive characteristics. They generally exhibit long excited-state lifetimes, high intersystem crossing efficiency, and good resistance to photobleaching and chemical degradation. Perhaps the most important feature of this class of compounds, however, is their structural modularity. The synthesis of heteroleptic $[\text{Ir}(\text{C}^{\wedge}\text{N})_2(\text{N}^{\wedge}\text{N})]^+$ complexes (where $\text{C}^{\wedge}\text{N}$ is a cyclometalating 2-phenylpyridinato ligand and $\text{N}^{\wedge}\text{N}$ is a diimine ancillary ligand) is straightforward and robust,¹³⁻²² and the excited-state properties of many structurally modified Ir(III) complexes have been reported.^{22,23} Thus, modification of the ligands on these compounds can provide a straightforward means to tune the energy of the emissive excited state for a given chemical application. The ability to *rationally* and *rapidly* predict the excited-state properties of novel Ir(III) chromophores would be a powerfully enabling capability in all the fields in which they have proven to be useful. Quantitative prediction of the absolute redox potentials of the singlet and the triplet states, the phosphorescence energies, and the emission spectra of these complexes, however, ~~is beyond the capabilities of~~ remains a challenge for conventional computational methods.

The emitting triplet excited state of this class of Ir(III) chromophores generally belong to one of three different types: 1) predominantly $\text{C}^{\wedge}\text{N}$ ligand-centered (^3LC) character, 2) predominantly charge-transfer (^3CT) character (metal and $\text{C}^{\wedge}\text{N}$ ligand to $\text{N}^{\wedge}\text{N}$ ligand charge transfer), or 3) featuring significant contributions from both.²⁴⁻²⁶ ~~The phosphorescence state of these complexes, i.e., the lowest triplet excited state (Kasha state), thus contains signatures from multiple orbital transitions.~~²⁷⁻³¹ ~~Hence,~~ Hence, achieving quantitatively ac-

curate *in silico* ~~modeling predictions~~ of the phosphorescence energies and spectra ~~of these complexes by quantum mechanical (QM) methods requires multiple~~ across a diversity of functionalized complexes, exhibiting differing classes of excited states, remains extremely challenge. While density functional theory (DFT) ~~and/or time-dependent DFT (TD-DFT)~~ may exhibit good performance for a particular class of complexes, the results are often strongly functional dependent, with no a priori clear or universal optimal choice. As such, ~~the highest accuracy is typically obtained with correlated methods, such as~~ domain-based local pair natural orbital coupled cluster theory (DLPNO-CCSD(T)), ~~and/or time-dependent DFT (TD-DFT) steps~~ with their associated computational costs.³²⁻³⁷ Even then, ~~the most accurate of these QM methods,~~ Δ DFT-DLPNO-CCSD(T), had a root-mean-squared deviation of ~ 5 kcal/mol compared to experimental values of phosphorescence energies,³² while more common (and computationally efficient) DFT methods exhibit characteristic errors that are much higher. This would roughly correspond to a prediction uncertainty window of ~ 40 -45 nm and higher around the 500 nm region of the visible spectrum.

Machine learning (ML) and deep learning (DL) models have attracted high interest recently for the rapid and accurate predictions of molecular properties.³⁸⁻⁵² Concerning the photophysical properties of Ir(III) complexes, Hatanaka and coworkers used a combination of DFT calculations and machine learning to develop quantitative models to classify 148 Ir(III) complexes based on their computed luminescence quantum yields.⁵³ Later, Rasulev and coworkers developed models to predict the excitation wavelengths for a set of 47 Ir(III) complexes,⁵⁴ developing quantitative structure-property ~~relationship~~ relationships between the complex properties and the emission wavelengths.

Recently, Bernhard and co-workers published the high-throughput synthesis and spectral evaluation of a much larger dataset of 1440 $[\text{Ir}(\text{C}^{\wedge}\text{N})_2(\text{N}^{\wedge}\text{N})]^+$ complexes.²³ This extensive spectral dataset, obtained under uniform experimental conditions, provides an opportunity to develop and train predictive ML models,⁵⁵ which can potentially be utilized to explore a larger chemical space of $[\text{Ir}(\text{C}^{\wedge}\text{N})_2(\text{N}^{\wedge}\text{N})]^+$ complexes for promising properties as well as pro-

vide guiding principles for future synthetic efforts. Bernhard and co-workers utilized their aforementioned Ir-complex data to extract simple, quantitative structure-emission energy relationships using linear combinations of energy shifts ascribed to individual ligands.²³ Although their approach presented a framework to predict the emission energies of heteroleptic Ir(III) complexes, it is not a general predictive method and is limited to only the specific ligands contained within the original high-throughput dataset.

Subsequently, Kulik and coworkers developed more generalizable DL models⁵⁵ for this same dataset using easy-to-interpret chemical information from both the molecular structures^{56,57} and the properties of the ligands, establishing relationships with the associated photophysical properties of the complex, such as the mean emission energies of phosphorescence, the excitation state lifetimes, and the emission spectral integral.⁵⁵ Such models are generalizable to novel ligands outside the training dataset, and thus to arbitrary heteroleptic complexes. Moreover, the authors also conduct uncertainty-controlled chemical exploration of hypothetical Ir(III) complexes to identify promising ligands, thus demonstrating the scope of accelerated discovery of Ir(III) complexes using DL models.

The seminal work of Kulik and coworkers clearly establishes the promise of DL methods for the prediction of photophysical properties for Ir complexes. However, several critical issues remain open for further exploration. In particular, the emission spectra of Ir complexes are often not simple, featureless, Gaussians (particularly for complexes of ³LC character) but are characterized by multiple emission maxima that may be due to multiple emissive states and/or vibronic progressions.²³ As such, the prediction of simple emission maximum (or related scalar quantities) does not necessarily provide a comprehensive picture of the emission. Furthermore, the use of experimental (vs. computational) training data raises important questions as to how to deal with finite and variable signal-to-noise in the measured emission spectra, which thus cannot be treated as typical “gold-standard” training data. Both of these issues – among others – are addressed in the present development.

It is notable that previous attempts to learn continuous spectra curves are far fewer in

number compared to efforts aimed at predicting scalar quantities like emission maxima and excitation wavelengths.^{51,54,55,58,59} Rinke and co-workers utilized various DL architectures to predict the full molecular excitation spectra for organic molecules from the QM7B^{60,61} and QM9⁶² datasets.⁶³ However, the models were trained on the *computationally* generated spectra of 132k molecules. They were then tested against the computed spectra of diastereomers of a subset of 10k molecules within the training set. On the other hand, Park and co-workers had trained DL models using experimental data for combinations of 11k organic chromophores in 369 solvents.⁶⁴ Both of these studies used diverse chemical data representations as inputs for their DL models. These included techniques ranging from a Coulomb matrix encoding of the molecular structure for a multilayer perceptron (MLP) neural network, to detailed connectivity information of the atoms for a tensor representation of a molecule for graph convolutional networks (GCN). Notably, Park and co-workers built models consisting of two GCNs: one for the chromophore and the other for the solvent. These were concatenated and finally passed through an MLP to derive the optical properties.

In order to address the challenges laid out above, we trained DL models to predict the emission spectra of Ir(III) complexes using an extensive experimental dataset. While training on experimental data, we achieve robustness to the presence of noisy, low-quality training spectra. We tested the predictive performance of our DL models in two use cases: for heteroleptic complexes formed by novel *combinations* of “known” C^N and N^N ligands in the dataset, and for complexes including novel *ligands* (which cannot be found within any complex in the training dataset). We find that our DL models are sufficient to capture the emission profiles of even highly featured spectra, thus providing a comprehensive picture of the emission energies. Consistent with prior results,⁵⁵ we found the models to be more accurate in the first case of predicting the spectra for novel combinations of “known” ligands and subsequently, we mapped the performance of ~~the second case~~ predicting for complexes with “unseen” ligands in the chemical space to identify regions of high and low model accuracy. Notably, the DL models produced a predictive performance for phosphorescence energies

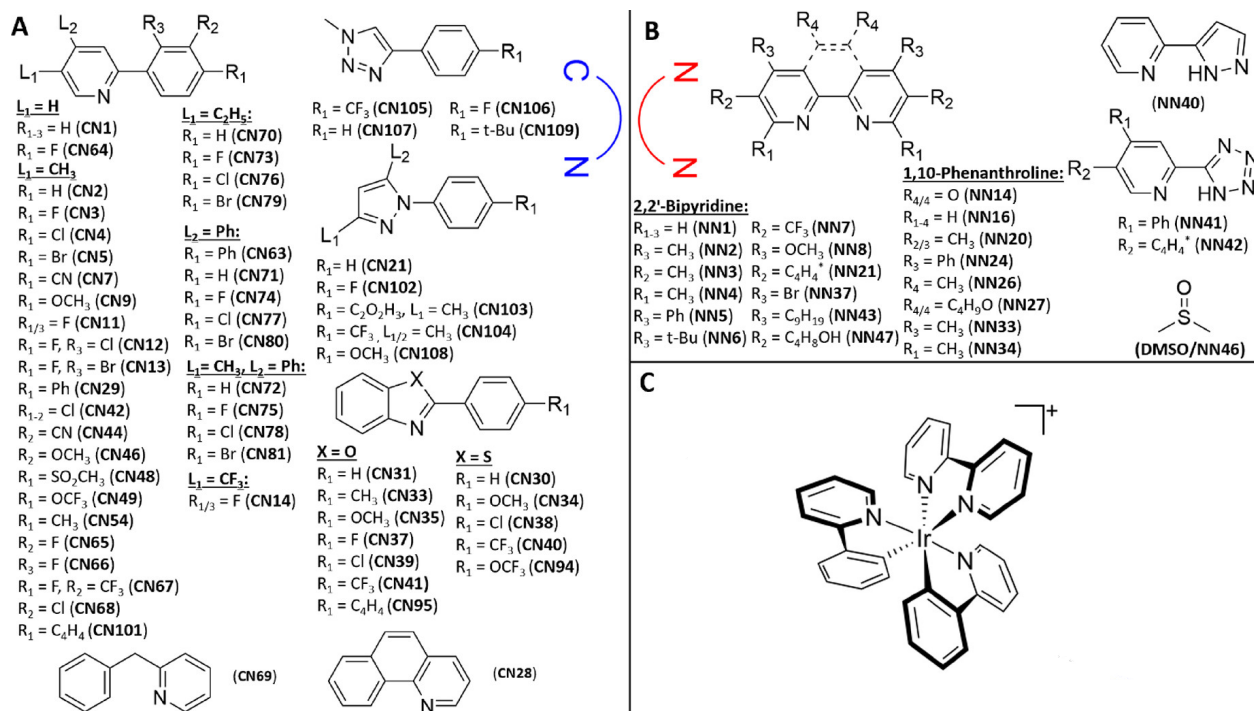
that significantly improved upon previously reported QM-based methods.

We further validated our DL model via additional experiments for several novel example complexes. For scenarios with poor blind predictive performance for novel ligands, we also found that having minimal additional data for these ligands greatly improves model accuracy and closes the gap in performance between the two use cases. These findings, described in the next sections, highlight our progress in obtaining general structure-luminescence relationships from the existing data toward the virtual discovery of new Ir(III) phosphorescent complexes, as well as the opportunities and challenges when leveraging experimental (vs. computational) data to train a data-driven model.

Methods

We have utilized the emission spectra data from the aforementioned work by Bernhard and co-workers²³ to train our models. The reference study’s complexes were produced by exhaustive combinations of 60 C[^]N and 24 N[^]N ligands, yielding 1440 total complexes. The 60 C[^]Ns include derivatives of several ~~2-phenylpyridinato~~-2-phenylpyridine archetypal molecules (Figure 1A), and the 24 N[^]Ns consisted of derivatives of 2,2’-~~bipyridine~~-bipyridine, 1,10-phenanthroline, tetrazoles, and pyrazoles (Figure 1B). Among the ligands, 2-phenylbenzothiazole- and 2-phenylbenzoxazole-derived C[^]Ns, and tetrazole- and pyrazole-derived N[^]Ns showed a higher tendency to exhibit emissive states of ³LC character.²³

It is important to note that since the same solvent (DMSO) and counterion (Cl⁻) were used consistently throughout the high-throughput synthesis, any potential solvatochromism or shifts due to ion interactions are implicitly included in the training data. That said, the solvatochromic shifts of this particular class of complexes are typically relatively modest,⁶⁵ while the high dielectric of the current solvent is expected to largely mitigate a potential ion pairing.⁶⁶ As such, the resulting model likely exhibits at least moderate transferability to other high-dielectric solvents.



$$w_i = \frac{1}{2}(1 + \tanh(\frac{I_i - 550}{100})) \quad (1)$$

A value of 550 counts was set as the origin, where the weighting function has a value of 0.5. As a result, most spectra were assigned weights approaching 1, while the spectra with reduced information content (such as those of complexes having a maximum absolute intensity below 200) carried smaller weights approaching 0, resulting in ~ 420 complexes out of 1440 having substantially reduced weights. The threshold value of 550 counts was determined qualitative examination of the spectra data. Briefly, the most intense emitters exhibit maximum counts in the thousands, whereas the weakest emitters (and background) is in the high tens of counts.²³ Spectra in the threshold regime exhibit well-defined peaks, but often exhibit clear noise on top of those peaks. This weighting process avoids the need for *ad hoc* culling of data from the training set and allows us to retain data that carries reduced (but non-zero) information content while also achieving robustness.

Our models were trained using a loss function consisting of the squared errors of the vectors of discretized, normalized spectral intensities over the fitted range (475 to 700 nm). However, because of the differential weighting of the training data, the cost function was defined as the weighted mean squared error (wMSE),

$$wMSE = \frac{\sum_{i=1}^{Test} (w_i * (\mathbf{y}_{pred,i} - \mathbf{y}_{true,i})^2)}{\sum_{j=1}^{Test} w_j} \quad (2)$$

We tested our models on randomized 80:20 training-testing splits of the exhaustive combinatorial dataset formed by the C^N and N^N ligands as well as ligand-specific splits. The predictive performance was evaluated by metrics such as the weighted root-mean-squared errors (wRMSE) and the weighted coefficients of determination (wR²) averaged across all 46 wavelengths,

$$wRMSE = \sqrt{\frac{\sum_{i=1}^{Test} (w_i * (\mathbf{y}_{pred,i} - \mathbf{y}_{true,i})^2)}{\sum_{j=1}^{Test} w_j}} \quad (3)$$

$$w\sigma^2 = \frac{\sum_{i=1}^{Test} (w_i * (\mathbf{y}_{true,i} - \frac{\sum_{j=1}^{Test} w_j * \mathbf{y}_{true,j}}{\sum_{k=1}^{Test} w_k})^2)}{\sum_{l=1}^{Test} w_l} \quad (4)$$

$$wR^2 = 1 - \frac{wMSE}{w\sigma^2} \quad (5)$$

We utilized 2 distinct frameworks for our DL models. The first of these involves a fully connected neural network i.e., MLP, with QM-calculated C^N and N^N ligand properties as the feature vectors. This simple structure naturally captures the diversity of the heteroleptic Ir complexes in terms of their constituent C^N and N^N ligand properties, yielding a relatively shallow neural network and a feature set of 16 scalar descriptors calculated for only 84 ligands. These features include intensive chemical properties of the ligands, such as the HOMO, LUMO, dipole moment, etc., which were chosen for their relevance to the ligand field theory-related phenomena involved in the emission of phosphorescent Ir complexes (a full list of the ligand properties used is shown in Table S1 in the SI). Most property features were obtained from singlet ground-state calculations. Some others like ionization energies, electron affinities, and singlet-triplet energies require multiple calculations and were chosen for their relevance to ligand-based charge transfer processes. We carried out DFT⁶⁷ calculations to obtain these ligand properties, using B3LYP-D3 theory⁶⁸⁻⁷¹ in combination with the 6-31G** basis set⁷² for all the atoms using Gaussian 16.⁷³ All DFT calculations included continuum solvation corrections with the dielectric constant of dimethyl sulfoxide (*via* IEF-PCM).⁷⁴⁻⁷⁶ Due to convergence issues with calculation of a small number of properties, we omitted a small number of C^N (6) and N^N (3) ligands from the final model.

We also examined an alternative graph network-based framework, which utilizes only the structure-connectivity information of the constituent C^N and N^N ligands (and thus no QM-calculated descriptors). This approach is based on MEGNet, a graph neural network (GNN)-based framework for universal DL for both molecular and crystalline materials.⁷⁷ In this framework, a molecule is represented as a graph with atoms at its vertices and bonds

at its edges. The attributes of each atom and bond, along with the state attributes for the whole molecule, are interconnected through the structure-connectivity framework of the molecular graph, thus encoding the chemical information of a molecule. Each “block” of MEGNet’s GNN conducts update operations on all the atom, bond, and state attributes by “pooling” information from the attributes of the respective nearest-neighbor bonds and atoms. Hence, successive blocks of MEGNet consider higher orders of nearest-neighbor interactions between the atoms and bonds as the molecular graph is processed through the GNN. MEGNet has been shown to outperform⁷⁷ other state-of-art models, such as SchNet⁷⁸ and MPNN models,⁷⁹ for predicting molecular properties. Thus, the MEGNet framework represents a more generalized and comprehensive approach compared to the shallow MLP framework described earlier.

To better reflect the structure of the heteroleptic Ir complexes, we generate a new graph network-based framework that initially encodes and processes the structural information of the C^N and N^N ligands in separate ~~graphs~~graph networks, before concatenating the output layers of the two graph networks together and ~~finally~~ running through an MLP to generate ~~the~~a vector of spectral intensities as output. We call this framework lig-MEGNet. The architectures of the 2 frameworks for our models discussed above are shown in Figure 2. lig-MEGNet contains 3 blocks of MEGNet convolution layers. The ligand-specific molecular graphs were obtained using MEGNet’s molecular graph converter module from the DFT-optimized XYZ files of the ligands. We also tested the lig-MEGNet architecture while simply using the ligands’ SMILES strings as input. The MLP and the lig-MEGNet-based models are similar in their architectures to models utilized by Kulik and coworkers.⁵⁵

The MLP models were developed with Keras⁸⁰ and trained using the Adam algorithm⁸¹ in mini-batches of 32 for 2000 epochs with an initial learning rate of 0.005 that was decayed by a factor of 0.5 down to a minimum of 0.0001. The lig-MEGNet models were trained using the Adam algorithm with the whole training set being treated as one batch and a learning rate of 0.01 for the first 500 epochs, 0.001 for the next 11000 epochs, and 0.0001 for the last

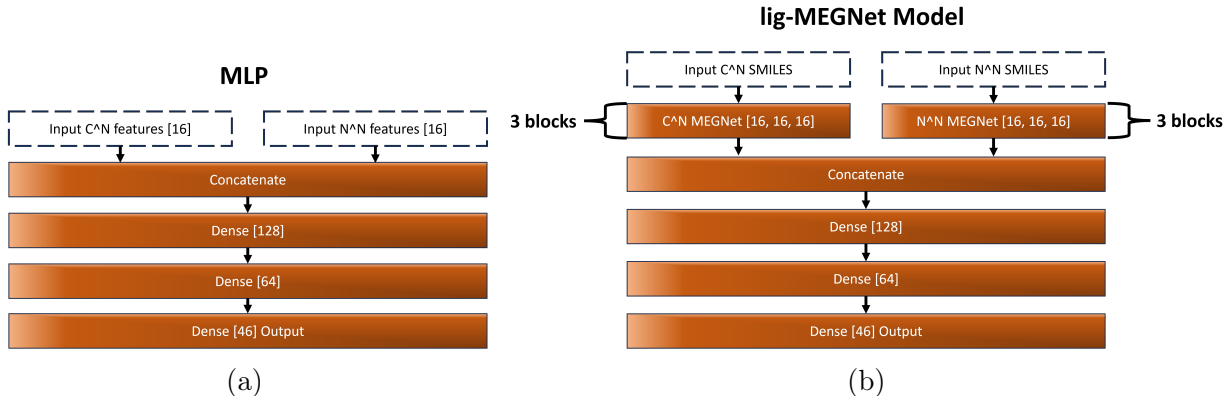


Figure 2: Architectures of the (a) MLP and b) lig-MEGNet frameworks. Inputs for the MLP models are vectors of 16 QM-calculated ligand property features for the C^N and N^N and for the lig-MEGNet models are molecular graphs with the bond, atom, and state attributes. The graphs ~~are~~ can be obtained from either the XYZ files of the C^N and the N^N ligands or simply from the SMILES strings of the ligands. See Ref.⁷⁷ for a more detailed explanation of the working of the GNN in a MEGNet “block,” ~~3 of which~~ MEGNet “blocks” were used in each ligand graph network in the lig-MEGNet framework. The “dense” layers are fully connected neural network layers. The final output layer for all models is a vector of 46 normalized intensities.

500 epochs. We chose these hyperparameters factoring in performance, speed, and memory usage after several trial runs.

Results and Discussion

Over tests on randomized 80:20 training-testing splits of the dataset, both models, i.e., MLP with ligand properties and lig-MEGNet, have comparable performance for predicting the spectral intensities, with test set wRMSEs of ~ 0.06 - 0.07 and wR^2 values that lie between 0.87 and 0.89 (see Table 1), when averaged across all 46 wavelengths. Although such GNNs have previously demonstrated strong performance on chemical problems, it is important to note that many of these applications were on vast and structurally diverse datasets, whereas the complexes in the current dataset exhibit strong structural similarity amongst the respective C^N and N^N classes, differing primarily *via* their differential functionalization in combinations of <100 unique ligands. Nonetheless, strong predictive performance is re-

tained. It is intriguing to note that the lig-MEGNet model is also able to achieve comparable performance even in the absence of any QM-calculated information (which includes chemically relevant descriptors, such as HOMO/LUMO energies, etc., or QM-optimized XYZ structural information), instead being driven by ligand structure-connectivity information through the SMILES strings. Note, however, that such GNNs may also require a larger dataset for training purposes as compared to a simple MLP based on ligand properties. This consideration could become relevant in regimes where training data is scarce.

Table 1: Performance of the different models for predicting spectral intensities and $\text{Em}_{50/50}$ over an 80:20 train-test split and the computational time required to converge the training process on a single NVIDIA 2080 GPU.

Model	Input data	Spectra Train wR^2	Spectra Test wR^2	Spectra Test wRMSE	$\text{Em}_{50/50}$ Test wR^2	$\text{Em}_{50/50}$ Test wRMSE (kcal/mol)	Time for training
MLP	QM ligand properties	0.97	0.89	0.06	0.93	0.51	2 min.
lig-MEGNet	ligand XYZ (QM)	0.97	0.87	0.07	0.90	0.56	10 min.
lig-MEGNet	ligand SMILES (no QM)	0.96	0.89	0.06	0.92	0.54	16 min.

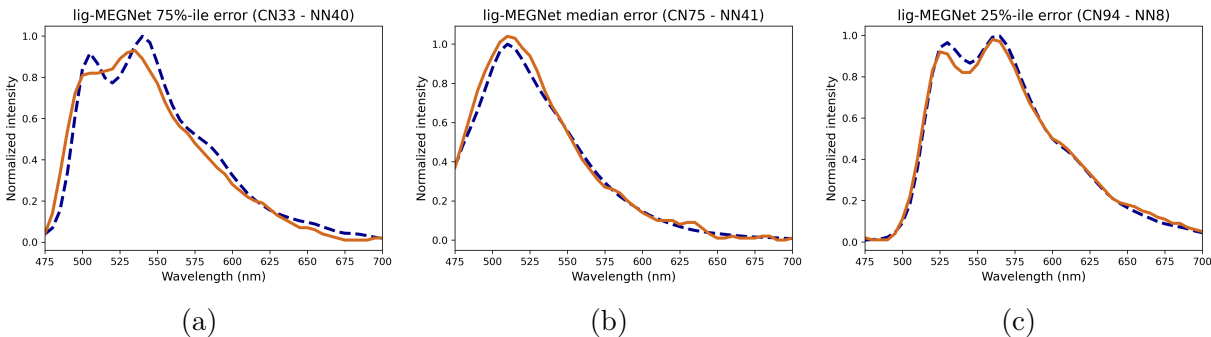


Figure 3: lig-MEGNet-predicted emission spectra (brown) *vs.* emission spectra from the high-throughput experiments of Bernhard and co-workers (dashed blue),²³ at the (a) 75th percentile, (b) median, and (c) 25th percentile points of test set error.

Figure 3 shows spectra predictions by the lig-MEGNet model (trained using QM-optimized XYZ geometries) for 3 complexes in the test set, at the 75th percentile, median, and 25th

percentile points of test set error. These samples are thus representative of the model’s predictive ability in various performance regimes. It is noteworthy that the model performs well in replicating spectra plots that display both the emission profile of ^3CT character (Figure 3b) as well as the vibronic emission profile of ^3LC character (Figure 3c). Plots of spectra predictions at the same 3 points of test set error by the MLP model can be found in Figures S2 and S3 of the SI. The corresponding $\text{Em}_{50/50}$ predictions have wRMSEs ~ 0.5 kcal/mol. Although, these models were trained to predict the emission spectra, the $\text{Em}_{50/50}$ predictions via the predicted spectra have similar errors to those shown by previous DL models developed on this dataset by Kulik and coworkers.⁵⁵ It is important to note that these errors, while not entirely trivial, are certainly smaller than can be achieved by even the state-of-the-art electronic structure theory methods on similar complexes (with, for example, a recent $\Delta\text{DFT-DLPNO-CCSD(T)}$ study exhibiting characteristic mean squared errors of ~ 5 kcal/mol).³² It is also possible that at least some of the residual error can likely be attributed to the inherent “uncertainty” of the high-throughput experimental training data: since the complexes were not isolated and purified prior to their characterization, there is likely some small (but unquantifiable) contribution from factors such as emission or adsorption from unreacted reactants or potential minor side products (*vide infra*).

The inclusion of ligand properties as state attributes to the lig-MEGNet framework, i.e., using both kinds of features, decreases the time needed to converge the training algorithm substantially (4 min *vs.* 10 min for lig-MEGNet, shown in Table S2) but does not lead to a significant improvement in performance. GNN models constructed using a single graph for the whole complex also do not yield much improvement, while requiring far greater training time. Thus, the design approach of representing the complexes by concatenating two separate graphs for the ligands naturally reflects the structural variance of the dataset and simultaneously optimizes the training requirements and the predictive performance.

Some of the complexes reported by Bernhard and co-workers²³ exhibit emission spectra with low signal-to-noise, possibly due to poor synthetic yields (e.g. ligands such as NN4,

NN7, NN14, N21, and NN37). The presence of such "noisy" emission spectra in the training set could result in substantial bias during the training process. As such, and as noted above, we thus employed a weighted training approach that reduces the training weight of complexes exhibiting weak emission intensity, but without requiring *ad hoc* culling of the training data. We compared a weighted model’s performance on the overall dataset with an unweighted model’s performance on a manually “cleaned” dataset (removing the ligands mentioned above), shown in Figure 4. The similar performance in both cases, with the highly weighted points close to the parity line and similar metrics, demonstrates the robustness of the training procedure even in the face of potentially low-fidelity training points.

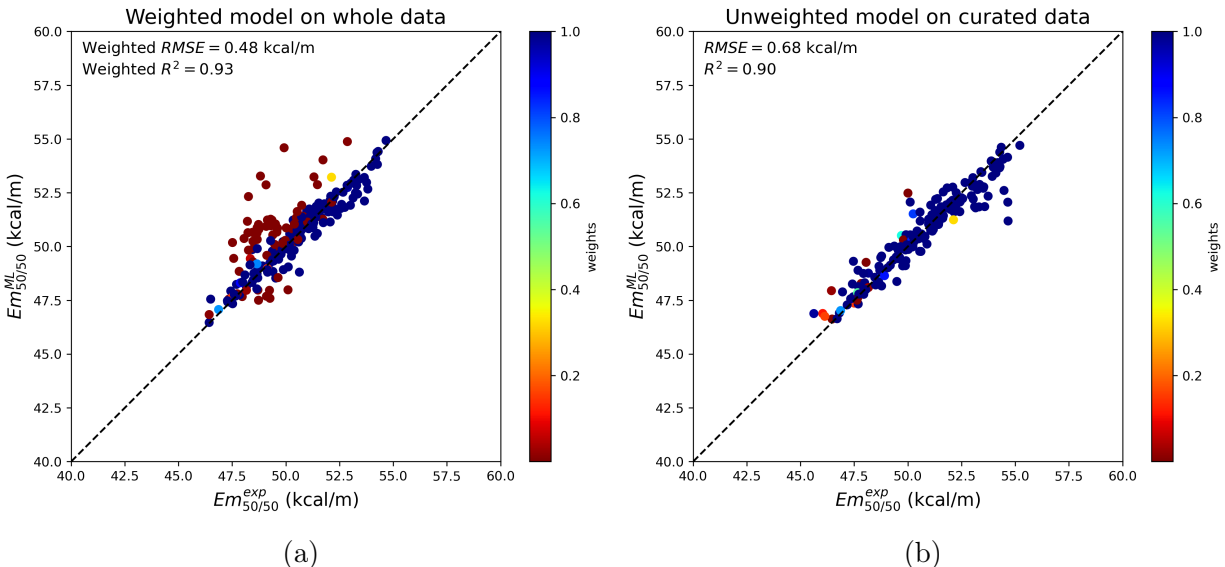


Figure 4: (a) Parity plot showing experimental values and predictions of $Em_{50/50}$ from a weighted model trained on the whole dataset. (b) Parity plot for an unweighted model trained on a curated dataset. The scattered points are shaded based on their weights.

We envision two potential use cases for our predictive models. One such use case would be to assist in the “design of experiment” in the sampling of a large space of heteroleptic complexes, such as may arise during a *partial* sampling of a high-throughput dataset similar to Bernhard and co-workers.²³ Although that dataset is exhaustive (it contained all potential combinations of C^N and N^N ligands within the sample space), one could envision a more efficient *partial* sampling of heteroleptic complexes that include only a small fraction of all

possible ligand combinations (while ensuring that all ligands arise in at least *some* of the sampled complexes). Here, a predictive ML/DL model would have great value in identifying promising ligand combinations within the larger sample space for further synthesis and characterization. Alternatively, one can envision applications to complexes that may contain entirely novel C^N or N^N ligands (or both), which have not been included in any of the training data. This case is expected to be more challenging since it necessitates extrapolation by the model in both the ligands space and the complexes space, whereas the prior application only requires extrapolation in the latter.

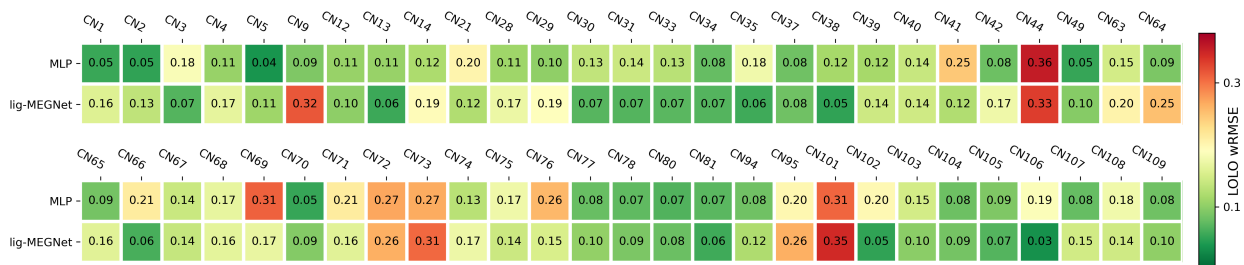


Figure 5: C^N ligand-wise LOLO wRMSEs of the predictions of spectral intensities by the MLP and the lig-MEGNet models.

Assessing the performance of our models in the second use case scenario, i.e., application to complexes containing novel ligands completely outside the original training set is more complicated. Rather than attempting to synthesize large quantities of complexes containing ligands outside the training set, we took a more pragmatic approach: we *simulated* this scenario by selective tailoring of the original training set *via* a “leave-one-ligand-out” (LOLO) approach. In this approach, we trained MLP and lig-MEGNet models over training sets that excluded *all* the complexes containing a *particular* C^N or an N^N ligand (blind LOLO). As such, the performance of a resulting model when applied to a complex that does contain the excluded ligand can be considered a truly blind prediction, identical to the scenario when a pre-trained model on the whole dataset is exposed to an “unknown” ligand in a new complex. The predictive performance over the C^N ligand-specific test sets is shown in Figure 5; a corresponding grid showing results over individual N^N ligands can be seen in Figure S4.

Examining the blind LOLO performance data, we find that it lags for C[^]N ligands belonging to archetypes that have fewer derived ligands relative to the 5-methyl-2-phenylpyridine-based groups, such as 5 ethyl-2-phenylpyridine (CN73, CN76), 2,4-diphenylpyridine (CN63, CN71, CN74), etc. MLP performs better than lig-MEGNet for complexes with ligands which typically exhibit emissive states of ³CT character in contrast to complexes with ligands exhibiting ³LC character (~~CNN30-CN41~~ CN30-CN41 and CN102-CN109 regions), where lig-MEGNet performs better. This is possibly due to the ligand features in the MLP being able to better capture the charge transfer processes in the more ³CT-character states than the intra-ligand interactions in the more ³LC-character states. Moreover, when we average over the performance of the 2 models across all the C[^]N ligands, we note that both models exhibit similar blind performances, with wRMSEs of spectral intensities of ~ 0.14 and wRMSEs of the corresponding Em_{50/50} of ~ 1.2 kcal/mol (Table 2). Similar to the performance difference between the grouped and random split testing cases in the earlier Em_{50/50} DL study by Kulik and coworkers,⁵⁵ these errors are higher than those observed in the random 80:20 split testing case (wRMSEs of ~ 0.06 for the intensities and ~ 0.5 kcal/mol for Em_{50/50}). Nonetheless, they still outperform the phosphorescence energy prediction errors of the state-of-the-art QM methods.

Table 2: Averaged predictive performance across all C[^]N ligands of the MLP and the lig-MEGNet models over test sets of complexes of individual ligands left out of the training sets (LOLO).

Model	Input data	Spectral intensities avg. wRMSE	Em _{50/50} avg. wRMSE (kcal/mol)
MLP	QM ligand properties	0.140	1.24
lig-MEGNet	ligand XYZ	0.139	1.20

Considering that these models utilize diverse levels of chemical information in networks of significantly different depths, their similar performance suggests that both models are close to saturation for learning from this dataset. Interestingly, when we compared the emission spectra from the high-throughput experiments of Bernhard and co-workers,²³ which were

used to train the models, and the emission spectra measured by us for a few of the complexes (see Figure S6 in the SI), we noticed a shift towards bluer wavelengths. One notable distinction between the two sets of experimental spectra data is that in the earlier experimental study by Bernhard and coworkers, the complexes were synthesized in a high-throughput manner, thus omitting product isolation and purification. We hypothesize that the presence of unreacted dimer complexes could slightly influence the measured emission spectra of the final product, which was observed for one concrete case of CN14-NN6 (Figure S6b and S6d). However, a quantitative correction is not possible due to uncertainties surrounding the reaction yield (and thus residual dimer concentration). Thus, considering this inherent residual uncertainty in the training dataset, we suspect that our models may have reached the limit of learning from the cost function over this dataset.

As expected, the performance of the blind LOLO lags that of the conventional randomized train/test splitting: in the LOLO analysis, the model is extrapolating in both the ligands space and the complexes space, since the model has (by construction) never seen any complex involving the ligand in question. From a practical point of view, it is natural to wonder what the *minimum* amount of training data/complexes is that would be required to reach performance comparable with the “conventional” scenario of randomized training-testing splits. This would dictate, for example, the number of complexes involving a given ligand that must be synthesized/characterized to allow for extrapolation from that data into the wider space of heteroleptic Ir(III) complexes. To address this question, we gradually added back between 1 and 11 complexes ([randomly chosen from the complexes](#) involving a candidate C^N ligand) into the formerly “blind” training set, resulting in a scenario intermediate between the two use cases described earlier. The resulting models were tested over that C^N ligand’s remaining complexes. Here, we present results only for the MLP models for brevity and efficiency.

Figure 6 shows results from this analysis for all C^N ligands. The top row of the grid shows the blind LOLO wRMSEs of the spectral intensities’ predictions for each of the C^Ns,

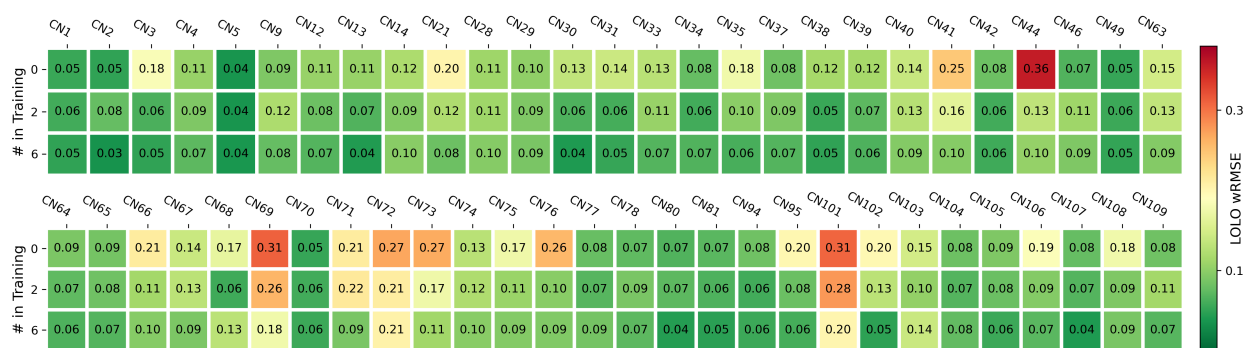


Figure 6: C^N ligand-wise LOLO wRMSEs of the predictions of spectral intensities *vs* the number of complexes of the C^N included in the training data.

with a ligand-average wRMSE of 0.14 (this is the same row as the top row from Figure 5). As expected, the performance of the models improves as they are exposed (in their training data) to complexes containing the test ligand, thus reducing the requirement for the models to extrapolate in the ligand space. However, what is striking is how *rapidly* the performance improves with only a tiny amount of additional data: adding only 2 complexes (containing the test C^N ligand) yields an average wRMSE of 0.10, a substantial improvement over the “blind” value of 0.14 and slightly higher than the wRMSE of 0.06 from the full training set of C^N ligands (Table 1).

The influence of additional data can be seen in Figure 7, where parity with training with a dataset containing the full set of C^Ns (Table 1) is reached by ~ 6 added complexes. Notably, we find significant inflection points in the curve when incorporating 1 or 2 complexes of the test C^N, suggesting diminishing gains in predictive accuracy as additional data is added. Critically, this suggests that future high-throughput screening studies could utilize similar ML/DL approaches to dramatically reduce the number of complexes that would need to be explicitly synthesized. Upon adding a novel C^N or N^N ligand, the synthesis of only a small number of additional complexes would be generally sufficient to predict the properties of a vast number of possible heteroleptic complexes involving that ligand.

It is interesting to note that while training on datasets with highly unbalanced numbers of complexes involving a given C^N ligand (such as the augmented LOLO analysis described

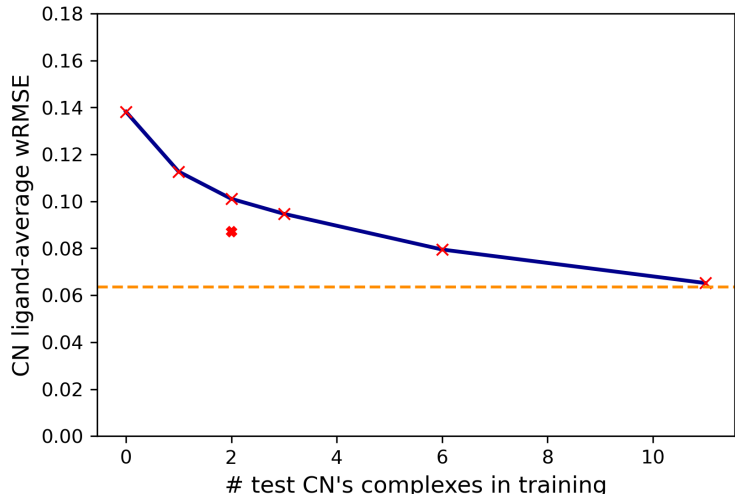


Figure 7: C^N ligand-averaged wRMSEs of the predictions of spectral intensities by the MLP LOLO models *vs* the number of complexes of a test C^N included in the training data (solid blue curve). The 0-complex point corresponds to the ligand average of the blind LOLO analysis (Table 2). The scenario of upweighting the data from a test C^N’s 2 complexes during training is represented by the bold red asterisk. The dashed orange line corresponds to the 80:20 train-test split wRMSE of the MLP model (Table 1) and serves as a reference for the asymptote of the plotted curve.

above), the imbalance may lead to correspondingly biased training and thus suboptimal utilization of the additional data. Therefore, we also tested a scenario where we up-weighted the data from the 2 included complexes of a test C^N by a factor of 10.5 ($=21/2$) during training to ensure a balanced representation, aligning the data of each of the other C^N’s 21 complexes (with the 21 N^Ns) at “parity” with the test C^N’s 2 complexes. In this case, we found the ligand-average wRMSE is further reduced to 0.09 (from 0.10, without up-weighting), and approaching even closer parity with the results from training with the full set of C^Ns.

It is important to note that the model’s success in predicting the blind, un-augmented LOLO analysis is far from uniform. For example, all models exhibit strong blind performance on CN34, despite never having encountered this ligand during training. In contrast, all models perform uniformly poorly on the blind prediction of CN101. We hypothesize that these variations are due to the extent to which the models are forced to extrapolate in the ligand space. To further shed light on the picture of ligand-wise variations in the models’

performance, we mapped them in the ligand properties space. We utilized two approaches for this. In the first, we compressed the C^N ligand properties space of 16 properties to a 2-component space through a t-distributed stochastic neighbor embedding (t-SNE) analysis.^{82,83} We generated a scatter plot of the C^N ligands in this 2-D space (Figure 8). In the second, we conducted a principal component (PC) analysis and projected the C^N ligands in a 2-D space of their 1st and 2nd PCs (Figure 9a) and their 1st and 3rd PCs (Figure 9b). The first 5 PCs account for 53.7%, 16.5%, 13.9%, 5.4%, and 3.6% of the total variance in the C^N properties data. Hence, in Figure 9, we exclusively mapped the 2nd and 3rd PCs against the 1st PC, which contribute significantly more to the variance compared to higher-numbered components. Both the t-SNE and the PCA analysis were carried out using Scikit-learn.⁸⁴ The t-SNE map helps us visualize similarities among various ligands by conserving relationships within localized data groupings. In contrast, the PC maps offer the added capability to visualize the *absolute* variance in the property features of the ligands across different linear projections. For both the t-SNE and the PC space scatter plots, the C^N ligand dots are colored based on their blind LOLO wRMSEs of intensities’ predictions.

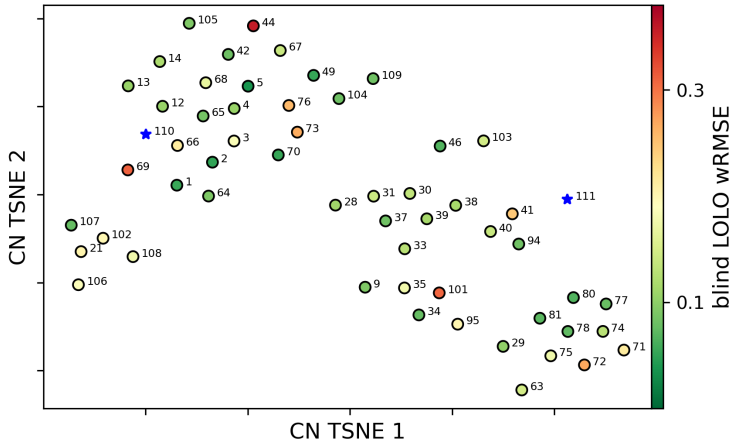


Figure 8: C^N ligands in a 2-D embedded space of their properties from a t-SNE analysis. The ligand dots are shaded by their blind LOLO wRMSEs of the MLP predictions of spectral intensities. The new C^Ns tested in Figures 10 and 11 are annotated as 110 and 111, respectively, and are represented as blue asterisks.

We note a general trend that similar ligands that cluster together in the 2-D t-SNE and

PC spaces tend to give lower wRMSEs for their blind LOLO predictions, whereas ligands with properties that are farther away from the neighborhood of other ligands tend to have higher wRMSEs. This trend is not just seen in the blind LOLO analysis but also in the augmented 2-in-training LOLO analysis (Figure S5). Especially from the mappings in the PC space (Figure 9), one can observe the dispersion in property values for ligands derived from 5-ethyl-2-phenylpyridine, 2,4-diphenylpyridine, and 5-methyl-2,4-diphenylpyridine (CN63, CN70-CN81), as well as 1-phenylpyrazole and 1-methyl-4-phenyl-1H-1,2,3-triazole (CN21, CN102-CN109). This dispersion is notable in contrast to a cluster of primarily 5-methyl-2-phenylpyridine-based ligands located near the center of the distribution. While the blind LOLO models are accurate for some of the distant ligands that form their own “mini-cluster” (CN77-CN81), C[^]Ns belonging to the archetypes mentioned above are more likely to yield poorer model performance.

To see how these insights from the LOLO, t-SNE, and PC analyses translate to novel complexes, we compared the models’ predictions for 2 example complexes against experimentally generated spectra. We synthesized these complexes with NN6 (4,4’-di-tert-butyl-2,2’-bipyridine) and two novel C[^]Ns (outside the original training set): one that closely resembles CN4, but with the Cl-group at a different position (CN110 in Figure 10); and another that is similar to CN14, but with an additional trifluoromethyl group (CN111 in Figure 11). The spectra predictions from the MLP and the lig-MEGNet models were plotted against our experimental spectra in Figures 10 and 11. The predictions are more accurate for CN110 (Figure 10) than for CN111. This is consistent with the fact that CN110 has more nearby neighbors in the t-SNE space (Figure 8) and is also closer to its neighbors in the PC space (Figure 9) when compared to CN111. Additionally, in the CN110 case, the lig-MEGNet-predicted spectra mimic the blueshifts seen between the high-throughput spectra of Bernhard and co-workers, which were used to train the models, and our measured spectra (see SI Section S9). As such, this shift may be at least partially due to the “bias” induced *via* training upon unpurified, high-throughput samples. Nevertheless, both models

predicted reasonably accurately for the CN110 ligand, which is from a data-rich region of the ligand space. The predicted spectra had errors lower than $\sim 15\text{-}20$ nm for the peak emission wavelength, which is similar to the uncertainty between the two sets of experiments (i.e., the prior high-throughput work²³ and those conducted by us as shown in Figure S6). In contrast, both models were similarly inaccurate for the CN111 ligand from a data-sparse region. Similar trends in predictive performance were also seen between the CN110-NN40 and CN111-NN40 complexes (Figure S7).

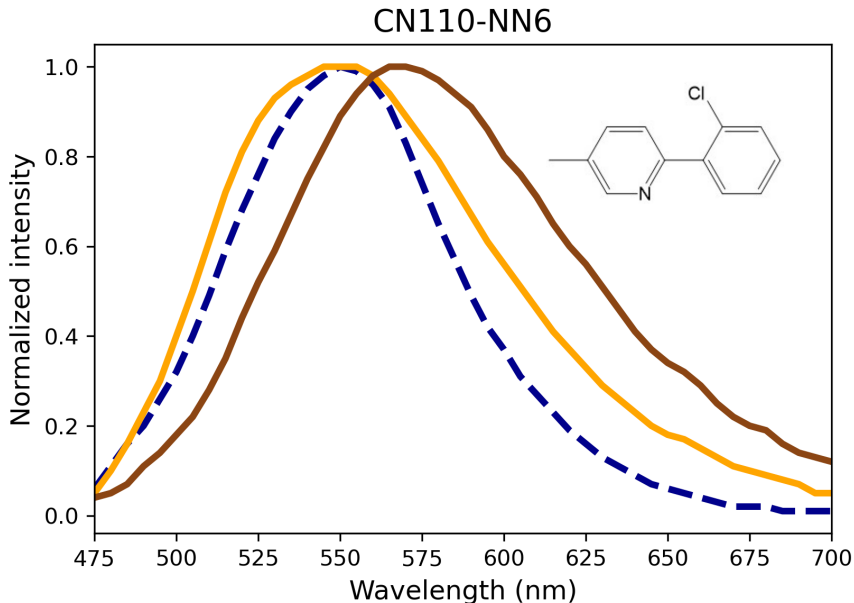


Figure 10: Predicted spectra using lig-MEGNet (brown) and MLP (orange) *vs* spectra measured in our laboratory (dashed blue) for the CN110-NN6 complex. Inset: CN110.

The results for CN111 further suggest that the models, in their current state, have limitations in extending their predictive capabilities to ligands representing substantial deviations in the chemical space from the training set. Even with the significantly more complex architecture of the MEGNet framework, these limitations apply to both the MLP and lig-MEGNet models. Nevertheless, as mentioned earlier, collecting even minimal additional experimental data in the sparsely populated regions of the ligand space could further enhance the accuracy of the models. Thus, the experimentally validated insights from the chemical space *vs* performance maps and the LOLO analyses showcase the utility of our findings for design-

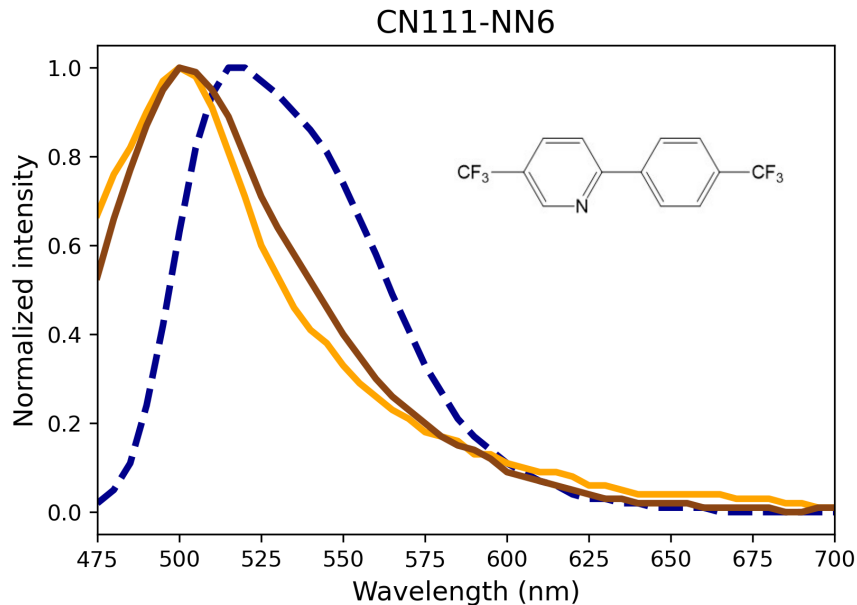


Figure 11: Predicted spectra using lig-MEGNet (brown) and MLP (orange) *vs* spectra measured in our laboratory (dashed blue) for the CN111-NN6 complex. Inset: CN111.

ing future experimental capacities for the rapid discovery of new phosphorescent complexes. Furthermore, we observed instances of ligands in data-sparse regions exhibiting low model uncertainty (Figure 9), indicating that models trained on existing data can also effectively identify new ligands and complexes of interest in certain cases. Overall, both potential use cases highlight how machine learning models can serve as a force multiplier by leveraging an initial set of training data to accelerate the discovery of optimal phosphorescent complexes.

Conclusion

We have presented a DL approach to predict the phosphorescence spectra of heteroleptic $[\text{Ir}(\text{C}^{\wedge}\text{N})_2(\text{N}^{\wedge}\text{N})]^+$ complexes using information from the molecular structures and/or properties of the constituent $\text{C}^{\wedge}\text{N}$ and $\text{N}^{\wedge}\text{N}$ ligands. Our models were trained directly on experimental data from ~ 1400 complexes, including appropriate weighting to account for the variable signal-to-noise of the underlying data, with the resulting models yielding smooth spectra predictions whose overall accuracies far surpass those of DFT and even state-of-

the-art correlated wavefunction methods. We found that DL architectures that reflect the inherent structure of the heteroleptic complexes, composed of C^N and N^N ligands, are both particularly effective and computationally efficient, yielding predicted emission energies *via* the predicted spectra to within ~ 1.2 kcal/mol of corresponding experimental measurements, similar to a recently published benchmark DL study by Kulik and coworkers.⁵⁵ Additionally our findings present a cautionary tale regarding the limitations in model accuracy due to the inherent uncertainties in the training data as well as in getting exact agreement between the predictions of a model trained on high-throughput experiments with standalone experiments. Nevertheless, we find that DL models of related structure exhibit strong potential for applications including “design of experiment” to mitigate synthetic effort in future high-throughput experimental studies.

However, we also find that the performance can be uneven, and may decline in regimes where the training set is data-poor in the “ligands space”, particularly when making blind predictions involving a complex with a previously “unseen” ligand. Although this observation would seem detrimental for applications of these and similar models in their most potent predictive capacity, we also find that this limitation can be largely mitigated by even a very modest amount of additional training data – as few as 1-2 additional complexes per new ligand. As such, the DL models retain significant promise as force multipliers, allowing the exploitation of limited amounts of experimental data (and thus synthetic effort) to explore a wide region of chemical space, and thus providing new pathways towards efficient design and optimization of phosphorescent materials.

Supporting Information Available

The Supporting Information is available free of charge at DOI:

- Experimental procedures; characterization data; ligand property features used; performance of additional model architectures; blind predictions over the set of N^N ligands;

comparison of predictions for additional complexes; supporting data and discussions.

- GitHub repository with code, data, and sample results can be accessed at https://github.com/yudhajitp/IrOLED_scripts_data

Acknowledgement

Funding for this work was provided by the 3M corporation and the NSF (CHE-195262). The NMR spectrometers used in this work are supported by the NSF CHE-1048642 and a generous gift from Paul J. and Margaret M. Bender. High Resolution Mass spectrometry was performed using a Thermo Q ExactiveTM Plus supported by the NIH 1S10 OD020022-1 and the University of Wisconsin.

References

- (1) Lowry, M. S.; Goldsmith, J. I.; Slinker, J. D.; Rohl, R.; Pascal, R. A.; Malliaras, G. G.; Bernhard, S. Single-Layer Electroluminescent Devices and Photoinduced Hydrogen Production from an Ionic Iridium(III) Complex. *Chemistry of Materials* **2005**, *17*, 5712–5719, doi: 10.1021/cm051312+.
- (2) Escudero, D. Quantitative prediction of photoluminescence quantum yields of phosphors from first principles. *Chem. Sci.* **2016**, *7*, 1262–1267.
- (3) Longhi, E.; Cola, L. D. *Iridium(III) Complexes for OLED Application*; John Wiley & Sons, Ltd, 2017; pp 205–274.
- (4) Jacquemin, D.; Escudero, D. The short device lifetimes of blue PhOLEDs: insights into the photostability of blue Ir(III) complexes. *Chem. Sci.* **2017**, *8*, 7844–7850.
- (5) Escudero, D. Mer-Ir(ppy)₃ to Fac-Ir(ppy)₃ Photoisomerization. *ChemPhotoChem* **2019**, *3*, 697–701.
- (6) L galit , F.; Escudero, D.; Pellegrin, Y.; Blart, E.; Jacquemin, D.; Fabrice, O. “Iridium effect” in cyclometalated iridium complexes for p-type dye sensitized solar cells. *Dyes and Pigments* **2019**, *171*, 107693.
- (7) Monos, T.; Stephenson, C. *Photoredox Catalysis of Iridium(III)-Based Photosensitizers*; 2017; pp 541–581.
- (8) Yano, N.; Handa, M.; Kataoka, Y. Photophysical properties and photosensitizing abilities for hydrogen evolution reactions of robust cyclometalated iridium(III) complexes with 5,5'-bis(trifluoromethyl)-2,2'-bipyridine. *Journal of Photochemistry and Photobiology A: Chemistry* **2020**, *400*, 112716.
- (9) Tritton, D. N.; Tang, F.-K.; Bodedla, G. B.; Lee, F.-W.; Kwan, C.-S.; Leung, K. C.-F.; Zhu, X.; Wong, W.-Y. Development and advancement of iridium(III)-based complexes

- for photocatalytic hydrogen evolution. *Coordination Chemistry Reviews* **2022**, *459*, 214390.
- (10) Zhang, L.; Ding, D. Recent advances of transition Ir(III) complexes as photosensitizers for improved photodynamic therapy. *VIEW* **2021**, *2*, 20200179.
- (11) Huang, H.; Banerjee, S.; Sadler, P. J. Recent Advances in the Design of Targeted Iridium(III) Photosensitizers for Photodynamic Therapy. *ChemBioChem* **2018**, *19*, 1574–1589.
- (12) Yu, M.; Zhao, Q.; Shi, L.; Li, F.; Zhou, Z.; Yang, H.; Yi, T.; Huang, C. Cationic iridium(iii) complexes for phosphorescence staining in the cytoplasm of living cells. *Chem. Commun.* **2008**, 2115–2117.
- (13) Tamayo, A. B.; Alleyne, B. D.; Djurovich, P. I.; Lamansky, S.; Tsyba, I.; Ho, N. N.; Bau, R.; Thompson, M. E. Synthesis and Characterization of Facial and Meridional Tris-cyclometalated Iridium(III) Complexes. *Journal of the American Chemical Society* **2003**, *125*, 7377–7387, doi: 10.1021/ja034537z.
- (14) Skubi, K. L.; Kidd, J. B.; Jung, H.; Guzei, I. A.; Baik, M.-H.; Yoon, T. P. Enantioselective Excited-State Photoreactions Controlled by a Chiral Hydrogen-Bonding Iridium Sensitizer. *Journal of the American Chemical Society* **2017**, *139*, 17186–17192, doi: 10.1021/jacs.7b10586.
- (15) Younker, J.; Dobbs, K. Correlating Experimental Photophysical Properties of Iridium(III) Complexes to Spin–Orbit Coupled TDDFT Predictions. *The Journal of Physical Chemistry C* **2013**, *117*, 25714–25723.
- (16) Singh, A.; Teegardin, K.; Kelly, M.; Prasad, K. S.; Krishnan, S.; Weaver, J. D. Facile synthesis and complete characterization of homoleptic and heteroleptic cyclometalated Iridium(III) complexes for photocatalysis. *Journal of Organometallic Chemistry* **2015**, *776*, 51–59.

- (17) Pomarico, E.; Silatani, M.; Messina, F.; Bräm, O.; Cannizzo, A.; Baranoff, E.; Klein, J.; Lambert, C.; Chergui, M. Dual Luminescence, Inter-Ligand Decay and Non-Radiative Electronic Relaxation of Cyclometalated Iridium Complexes in Solution. *The Journal of Physical Chemistry C* **2016**, *120*.
- (18) Deaton, J.; Castellano, F. *Archetypal Iridium(III) Compounds for Optoelectronic and Photonic Applications: Photophysical Properties and Synthetic Methods*; 2017; pp 1–69.
- (19) Kim, J.-H.; Kim, S.-Y.; Cho, Y.-J.; Son, H.-J.; Cho, D. W.; Kang, S. A Detailed Evaluation for the Nonradiative Processes in Highly Phosphorescent Iridium(III) Complexes. *The Journal of Physical Chemistry C* **2018**, *122*.
- (20) Soriano-Díaz, I.; Ortí, E.; Giussani, A. On the Importance of Ligand-Centered Excited States in the Emission of Cyclometalated Ir(III) Complexes. *Inorganic Chemistry* **2021**, *60*, 13222–13232, doi: 10.1021/acs.inorgchem.1c01604.
- (21) Florent, M.; Kyritsakas, N.; Planeix, J.-M.; Guenet, A.; Hosseini, M. W. Luminescent 1D heterometallic (Ir,Cd) coordination polymers based on bis-cyclometalated Ir(iii) metallatectons and trinuclear Cd(ii) dianionic nodes. *Dalton Trans.* **2021**, *50*, 15924–15934.
- (22) Bawden, J. C.; Francis, P. S.; DiLuzio, S.; Hayne, D. J.; Doeven, E. H.; Truong, J.; Alexander, R.; Henderson, L. C.; Gómez, D. E.; Massi, M. et al. Reinterpreting the Fate of Iridium(III) Photocatalysts-Screening a Combinatorial Library to Explore Light-Driven Side-Reactions. *Journal of the American Chemical Society* **2022**, *144*, 11189–11202, doi: 10.1021/jacs.2c02011.
- (23) DiLuzio, S.; Mdluli, V.; Connell, T. U.; Lewis, J.; VanBenschoten, V.; Bernhard, S. High-Throughput Screening and Automated Data-Driven Analysis of the Triplet Photophysical Properties of Structurally Diverse, Heteroleptic Iridium(III) Complexes. *Journal of the American Chemical Society* **2021**, *143*, 1179–1194, doi: 10.1021/jacs.0c12290.

- (24) Colombo, M. G.; Hauser, A.; Gudel, H. U. Evidence for strong mixing between the LC and MLCT excited states in bis(2-phenylpyridinato-C2,N')(2,2'-bipyridine)iridium(III). *Inorganic Chemistry* **1993**, *32*, 3088–3092, doi: 10.1021/ic00066a020.
- (25) Yang, C.-H.; Li, S.-W.; Chi, Y.; Cheng, Y.-M.; Yeh, Y.-S.; Chou, P.-T.; Lee, G.-H.; Wang, C.-H.; Shu, C.-F. Heteroleptic Cyclometalated Iridium(III) Complexes Displaying Blue Phosphorescence in Solution and Solid State at Room Temperature. *Inorganic Chemistry* **2005**, *44*, 7770–7780, doi: 10.1021/ic050311g.
- (26) Tamayo, A. B.; Garon, S.; Sajoto, T.; Djurovich, P. I.; Tsyba, I. M.; Bau, R.; Thompson, M. E. Cationic Bis-cyclometalated Iridium(III) Diimine Complexes and Their Use in Efficient Blue, Green, and Red Electroluminescent Devices. *Inorganic Chemistry* **2005**, *44*, 8723–8732, doi: 10.1021/ic050970t.
- (27) Escudero, D.; Jacquemin, D. Computational insights into the photodeactivation dynamics of phosphors for OLEDs: a perspective. *Dalton Trans.* **2015**, *44*, 8346–8355.
- (28) Yersin, H.; Rausch, A. F.; Czerwieniec, R.; Hofbeck, T.; Fischer, T. The triplet state of organo-transition metal compounds. Triplet harvesting and singlet harvesting for efficient OLEDs. *Coordination Chemistry Reviews* **2011**, *255*, 2622–2652, Controlling photophysical properties of metal complexes: Towards molecular photonics.
- (29) Zhang, K. Y.; Gao, P.; Sun, G.; Zhang, T.; Li, X.; Liu, S.; Zhao, Q.; Lo, K. K.-W.; Huang, W. Dual-Phosphorescent Iridium(III) Complexes Extending Oxygen Sensing from Hypoxia to Hyperoxia. *Journal of the American Chemical Society* **2018**, *140*, 7827–7834, doi: 10.1021/jacs.8b02492.
- (30) Kumar, S.; Hisamatsu, Y.; Tamaki, Y.; Ishitani, O.; Aoki, S. Design and Synthesis of Heteroleptic Cyclometalated Iridium(III) Complexes Containing Quinoline-Type Ligands that Exhibit Dual Phosphorescence. *Inorganic Chemistry* **2016**, *55*, 3829–3843, doi: 10.1021/acs.inorgchem.5b02872.

- (31) Yeh, Y.-S.; Cheng, Y.-M.; Chou, P.-T.; Lee, G.-H.; Yang, C.-H.; Chi, Y.; Shu, C.-F.; Wang, C.-H. A New Family of Homoleptic Ir(III) Complexes: Tris-Pyridyl Azolate Derivatives with Dual Phosphorescence. *ChemPhysChem* **2006**, *7*, 2294–2297.
- (32) Kumar, P.; Pérez-Escribano, M.; van Raamsdonk, D. M. E.; Escudero, D. Phosphorescent Properties of Heteroleptic Ir(III) Complexes: Uncovering Their Emissive Species. *The Journal of Physical Chemistry A* **2023**, *127*, 7241–7255, doi: 10.1021/acs.jpca.3c04205.
- (33) Kundu, S.; Roy, P. P.; Fleming, G. R.; Makri, N. Franck–Condon and Herzberg–Teller Signatures in Molecular Absorption and Emission Spectra. *The Journal of Physical Chemistry B* **2022**, *126*, 2899–2911, doi: 10.1021/acs.jpcb.2c00846.
- (34) Baiardi, A.; Bloino, J.; Barone, V. General Time Dependent Approach to Vibronic Spectroscopy Including Franck–Condon, Herzberg–Teller, and Duschinsky Effects. *Journal of Chemical Theory and Computation* **2013**, *9*, 4097–4115, doi: 10.1021/ct400450k.
- (35) Roy, P. P.; Kundu, S.; Makri, N.; Fleming, G. R. Interference between Franck–Condon and Herzberg–Teller Terms in the Condensed-Phase Molecular Spectra of Metal-Based Tetrapyrrole Derivatives. *The Journal of Physical Chemistry Letters* **2022**, *13*, 7413–7419, doi: 10.1021/acs.jpcllett.2c01963.
- (36) Grotjahn, R.; Kaupp, M. Reliable TDDFT Protocol Based on a Local Hybrid Functional for the Prediction of Vibronic Phosphorescence Spectra Applied to Tris(2,2'-bipyridine)-Metal Complexes. *The Journal of Physical Chemistry A* **2021**, *125*, 7099–7110, doi: 10.1021/acs.jpca.1c05101.
- (37) Mori, K.; Goumans, T. P. M.; van Lenthe, E.; Wang, F. 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.

- (38) Behler, J. Four Generations of High-Dimensional Neural Network Potentials. *Chemical Reviews* **2021**, *121*, 10037–10072, doi: 10.1021/acs.chemrev.0c00868.
- (39) Deringer, V. L.; Bartók, A. P.; Bernstein, N.; Wilkins, D. M.; Ceriotti, M.; Csányi, G. Gaussian Process Regression for Materials and Molecules. *Chemical Reviews* **2021**, *121*, 10073–10141, doi: 10.1021/acs.chemrev.1c00022.
- (40) Fedik, N.; Zubatyuk, R.; Kulichenko, M.; Lubbers, N.; Smith, J. S.; Nebgen, B.; Messerly, R.; Li, Y. W.; Boldyrev, A. I.; Barros, K. et al. Extending machine learning beyond interatomic potentials for predicting molecular properties. *Nature Reviews Chemistry* **2022**, *6*, 653–672.
- (41) Staszak, M.; Staszak, K.; Wieszczycka, K.; Bajek, A.; Roszkowski, K.; Tylkowski, B. Machine learning in drug design: Use of artificial intelligence to explore the chemical structure–biological activity relationship. *WIREs Computational Molecular Science* **2022**, *12*, e1568.
- (42) Reiser, P.; Neubert, M.; Eberhard, A.; Torresi, L.; Zhou, C.; Shao, C.; Metni, H.; van Hoesel, C.; Schopmans, H.; Sommer, T. et al. Graph neural networks for materials science and chemistry. *Communications Materials* **2022**, *3*, 93.
- (43) Walters, W. P.; Barzilay, R. Applications of Deep Learning in Molecule Generation and Molecular Property Prediction. *Accounts of Chemical Research* **2021**, *54*, 263–270, doi: 10.1021/acs.accounts.0c00699.
- (44) Reiser, P.; Konrad, M.; Fediai, A.; Léon, S.; Wenzel, W.; Friederich, P. Analyzing Dynamical Disorder for Charge Transport in Organic Semiconductors via Machine Learning. *Journal of Chemical Theory and Computation* **2021**, *17*, 3750–3759, doi: 10.1021/acs.jctc.1c00191.
- (45) Morawietz, T.; Singraber, A.; Dellago, C.; Behler, J. How van der Waals interactions

- determine the unique properties of water. *Proceedings of the National Academy of Sciences* **2016**, *113*, 8368–8373.
- (46) Deringer, V. L.; Bernstein, N.; Csányi, G.; Mahmoud, C. B.; Ceriotti, M.; Wilson, M.; Drabold, D. A.; Elliott, S. R. Origins of structural and electronic transitions in disordered silicon. *Nature* **2021**, *589*, 59–64.
- (47) Cheng, B.; Mazzola, G.; Pickard, C. J.; Ceriotti, M. Evidence for supercritical behaviour of high-pressure liquid hydrogen. *Nature* **2020**, *585*, 217–220.
- (48) Kapil, V.; Schran, C.; Zen, A.; Chen, J.; Pickard, C. J.; Michaelides, A. The first-principles phase diagram of monolayer nanoconfined water. *Nature* **2022**, *609*, 512–516.
- (49) Stocker, S.; Csányi, G.; Reuter, K.; Margraf, J. T. Machine learning in chemical reaction space. *Nature Communications* **2020**, *11*, 5505.
- (50) Gómez-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. *Nature Materials* **2016**, *15*, 1120–1127.
- (51) Ju, C.-W.; Bai, H.; Li, B.; Liu, R. Machine Learning Enables Highly Accurate Predictions of Photophysical Properties of Organic Fluorescent Materials: Emission Wavelengths and Quantum Yields. *Journal of Chemical Information and Modeling* **2021**, *61*, 1053–1065, doi: 10.1021/acs.jcim.0c01203.
- (52) Sifain, A. E.; Lystrom, L.; Messerly, R. A.; Smith, J. S.; Nebgen, B.; Barros, K.; Tretiak, S.; Lubbers, N.; Gifford, B. J. Predicting phosphorescence energies and inferring wavefunction localization with machine learning. *Chem. Sci.* **2021**, *12*, 10207–10217.
- (53) Hatanaka, M.; Kato, H.; Sakai, M.; Kariya, K.; Nakatani, S.; Yoshimura, T.; Inagaki, T. Insights into the Luminescence Quantum Yields of Cyclometalated Iridium(III) Com-

- plexes: A Density Functional Theory and Machine Learning Approach. *The Journal of Physical Chemistry A* **2023**, *127*, 7630–7637, doi: 10.1021/acs.jpca.3c02179.
- (54) Karuth, A.; Casanola-Martin, G. M.; Lystrom, L.; Sun, W.; Kilin, D.; Kilina, S.; Rasulev, B. Combined Machine Learning, Computational, and Experimental Analysis of the Iridium(III) Complexes with Red to Near-Infrared Emission. *The Journal of Physical Chemistry Letters* **2024**, *15*, 471–480, doi: 10.1021/acs.jpclett.3c02533.
- (55) Terrones, G. G.; Duan, C.; Nandy, A.; Kulik, H. J. Low-cost machine learning prediction of excited state properties of iridium-centered phosphors. *Chem. Sci.* **2023**, *14*, 1419–1433.
- (56) Janet, J. P.; Kulik, H. J. Resolving Transition Metal Chemical Space: Feature Selection for Machine Learning and Structure–Property Relationships. *The Journal of Physical Chemistry A* **2017**, *121*, 8939–8954, doi: 10.1021/acs.jpca.7b08750.
- (57) Duan, C.; Liu, F.; Nandy, A.; Kulik, H. J. Data-Driven Approaches Can Overcome the Cost–Accuracy Trade-Off in Multireference Diagnostics. *Journal of Chemical Theory and Computation* **2020**, *16*, 4373–4387, doi: 10.1021/acs.jctc.0c00358.
- (58) Kang, B.; Seok, C.; Lee, J. Prediction of Molecular Electronic Transitions Using Random Forests. *Journal of Chemical Information and Modeling* **2020**, *60*, 5984–5994, doi: 10.1021/acs.jcim.0c00698.
- (59) Ye, Z.-R.; Huang, I.-S.; Chan, Y.-T.; Li, Z.-J.; Liao, C.-C.; Tsai, H.-R.; Hsieh, M.-C.; Chang, C.-C.; Tsai, M.-K. Predicting the emission wavelength of organic molecules using a combinatorial QSAR and machine learning approach. *RSC Adv.* **2020**, *10*, 23834–23841.
- (60) Blum, L. C.; Raymond, J.-L. 970 Million Druglike Small Molecules for Virtual Screening in the Chemical Universe Database GDB-13. *Journal of the American Chemical Society* **2009**, *131*, 8732–8733, doi: 10.1021/ja902302h.

- (61) Montavon, G.; Rupp, M.; Gobre, V.; Vazquez-Mayagoitia, A.; Hansen, K.; Tkatchenko, A.; Müller, K.-R.; von Lilienfeld, O. A. Machine learning of molecular electronic properties in chemical compound space. *New Journal of Physics* **2013**, *15*, 95003.
- (62) Ramakrishnan, R.; Dral, P. O.; Rupp, M.; von Lilienfeld, O. A. Quantum chemistry structures and properties of 134 kilo molecules. *Scientific Data* **2014**, *1*, 140022.
- (63) Ghosh, K.; Stuke, A.; Todorović, M.; Jørgensen, P. B.; Schmidt, M. N.; Vehtari, A.; Rinke, P. Deep Learning Spectroscopy: Neural Networks for Molecular Excitation Spectra. *Advanced Science* **2019**, *6*, 1801367.
- (64) Joung, J. F.; Han, M.; Hwang, J.; Jeong, M.; Choi, D. H.; Park, S. Deep Learning Optical Spectroscopy Based on Experimental Database: Potential Applications to Molecular Design. *JACS Au* **2021**, *1*, 427–438, doi: 10.1021/jacsau.1c00035.
- (65) Bryden, M. A.; Millward, F.; Lee, O. S.; Cork, L.; Gather, M. C.; Steffen, A.; Zysman-Colman, E. Lessons learnt in photocatalysis – the influence of solvent polarity and the photostability of the photocatalyst. *Chem. Sci.* **2024**, –.
- (66) Earley, J. D.; Zieleniewska, A.; Ripberger, H. H.; Shin, N. Y.; Lazorski, M. S.; Mast, Z. J.; Sayre, H. J.; McCusker, J. K.; Scholes, G. D.; Knowles, R. R. et al. Ion-pair reorganization regulates reactivity in photoredox catalysts. *Nature Chemistry* **2022**, *14*, 746–753.
- (67) Parr, R. G. Density Functional Theory of Atoms and Molecules. 1980; pp 5–15.
- (68) 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.
- (69) Becke, A. D. Density-functional thermochemistry. I. The effect of the exchange-only gradient correction. *The Journal of Chemical Physics* **1992**, *96*, 2155–2160.

- (70) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *The Journal of Chemical Physics* **1993**, *98*, 5648–5652.
- (71) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *The Journal of Chemical Physics* **2010**, *132*, 154104.
- (72) Ditchfield, R.; Hehre, W. J.; Pople, J. A. Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules. *The Journal of Chemical Physics* **2003**, *54*, 724–728.
- (73) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H. et al. Gaussian~16 Revision C.01. 2016; Gaussian Inc. Wallingford CT.
- (74) Miertuš, S.; Scrocco, E.; Tomasi, J. Electrostatic interaction of a solute with a continuum. A direct utilization of AB initio molecular potentials for the prevision of solvent effects. *Chemical Physics* **1981**, *55*, 117–129.
- (75) Miertuš, S.; Tomasi, J. Approximate evaluations of the electrostatic free energy and internal energy changes in solution processes. *Chemical Physics* **1982**, *65*, 239–245.
- (76) Pascual-ahuir, J. L.; Silla, E.; Tuñon, I. GEPOL: An improved description of molecular surfaces. III. A new algorithm for the computation of a solvent-excluding surface. *Journal of Computational Chemistry* **1994**, *15*, 1127–1138.
- (77) Chen, C.; Ye, W.; Zuo, Y.; Zheng, C.; Ong, S. P. Graph Networks as a Universal Machine Learning Framework for Molecules and Crystals. *Chemistry of Materials* **2019**, *31*, 3564–3572, doi: 10.1021/acs.chemmater.9b01294.
- (78) Schütt, K. T.; Sauceda, H. E.; Kindermans, P.-J.; Tkatchenko, A.; Müller, K.-R. SchNet

- A deep learning architecture for molecules and materials. *The Journal of Chemical Physics* **2018**, *148*, 241722.
- (79) Gilmer, J.; Schoenholz, S. S.; Riley, P. F.; Vinyals, O.; Dahl, G. E. Neural Message Passing for Quantum Chemistry. 2017; Accessed: October 10, 2023.
- (80) Chollet, F. Keras. <https://keras.io/>, 2015; Accessed: October 10, 2023.
- (81) Kingma, D. P.; Ba, J. Adam: A Method for Stochastic Optimization. 2017; Accessed: October 10, 2023.
- (82) Hinton, G. E.; Roweis, S. Stochastic Neighbor Embedding. 2002.
- (83) van der Maaten, L.; Hinton, G. Visualizing Data using t-SNE. *Journal of Machine Learning Research* **2008**, *9*, 2579–2605.
- (84) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V. et al. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **2011**, *12*, 2825–2830.