Stacked Ensemble Machine Learning for Range-Separation Parameters

Cheng-Wei Ju,†,‡ Ethan J. French,† Nadav Geva,¶ Alexander W. Kohn,§ and Zhou Lin*,†

†Department of Chemistry, University of Massachusetts, Amherst, Massachusetts 01003, United States
‡College of Chemistry, Nankai University, Tianjin 300071, China
¶Advanced Micro Devices Inc., Boxborough, Massachusetts 01719, United States
§Blizzard Entertainment Inc., Irvine, California 92618, United States

E-mail: zhoulin@umass.edu
Abstract

High-throughput virtual materials and drug discovery based on density functional theory has achieved tremendous success in recent decades, but its power on organic semiconducting molecules suffers catastrophically from self-interaction error until the optimally tuned range-separated hybrid (OT-RSH) exchange–correlation functionals were developed. The accurate but expensive first-principles OT-RSH transitions from a short-range (semi-)local functional to a long-range Hartree–Fock exchange at a distance characterized by the inverse of a molecule-specific, non-empirically-determined range-separation parameter ($\omega$). In the present study, we proposed a promising stacked ensemble machine learning model that provides an accelerated alternative of OT-RSH based on system-dependent structural and electronic configurations. We trained ML-$\omega$PBE, the first functional in our series, using a database of 1,970 organic semiconducting molecules with sufficient structural diversity, and assessed its accuracy and efficiency using another 1,956 molecules. Compared with the first-principles OT-$\omega$PBE, our ML-$\omega$PBE reached a mean absolute error of $0.00504a_0^{-1}$ for the optimal value of $\omega$, reduced the computational cost for the test set by 2.66 orders of magnitude, and achieved a comparable predictive power in various optical properties.
Organic semiconducting molecules with large-scale π-conjugations\(^1\)–\(^3\) have been the spotlight of next-generation materials\(^4\)–\(^8\) and biological sciences\(^9\)–\(^13\) since the 1960s because of their compelling electronic and optical properties. Due to the ever-growing demand for materials and drug discovery, high-throughput development of organic semiconducting molecules gradually shift gear from old-school trial-and-error experiments to first-principles approaches like density functional theory (DFT).\(^14\)–\(^19\) One outstanding example for such a computational design was thermally activated delayed fluorescence (TADF) emitters, for which overall fluorescence quantum yields were predicted based on the DFT-evaluated electronic configurations, especially the alignment of frontier orbitals.\(^20\)–\(^22\)

The reliability of DFT can however be impaired by the self-interaction error (SIE)\(^23\),\(^24\) of a (semi-)local exchange–correlation (XC) functional, which leads to an overdelocalized electronic density and an incorrect derivative discontinuity (DD),\(^25\)–\(^30\) and eventually a violation of Koopmans’\(^31\) and Janak’s\(^32\) theorems and an underestimated charge transfer excitation energy.\(^33\)–\(^36\) An effective solution is an error-canceling hybrid of the exact Hartree–Fock (HF) exchange energy and some (semi-)local XC functionals, including global hybrids,\(^37\)–\(^41\) as well as range-separated hybrids (RSHs) which typically utilize the scheme of\(^34\),\(^42\)–\(^66\)

\[
\frac{1}{|\mathbf{r} - \mathbf{r}'|} = \frac{1 - \text{erf} (\omega |\mathbf{r} - \mathbf{r}'|)}{|\mathbf{r} - \mathbf{r}'|} + \frac{\text{erf} (\omega |\mathbf{r} - \mathbf{r}'|)}{|\mathbf{r} - \mathbf{r}'|}
\]

(1)

Here \(r_{12}\) represents the interelectron distance and erf the Gauss error function. The range-separation parameter (\(\omega\)) characterizes the inverse of distance where the short-range (SR) (semi-)local functional transitions into the long-range (LR) HF exchange. Its ideal value is sensitive to the structural and electronic configuration of a molecule and can be optimally tuned (OT) based on Koopmans’ theorem\(^31\) by minimizing

\[
J^2 (\omega) = [\varepsilon_{\text{HOMO}} (\omega) + I (\omega)]^2 + [\varepsilon_{\text{LUMO}} (\omega) + A (\omega)]^2
\]

(2)
where HOMO and LUMO represent the highest occupied and lowest unoccupied molecular orbitals, and \( I \) and \( A \) describe the ionization potential and the electron affinity, respectively.

OT-RSH functionals possessed promoted predictive powers in electronic and optical properties properties,\(^{33,67–81}\) but they are expensive because the optimization of \( J^2(\omega) \) typically requires 20 to 50 converged self-consistent field calculations in the lack of an analytical gradient. Multiple machine learning (ML) strategies have been reported to alleviate the problem.\(^{82–95}\) For instance, Chen and coworkers trained long-range-corrected Becke–Lee–Yang–Parr functional using the neural network as the algorithm (LC-BLYP-NN) and atom–atom interactions as molecular descriptors, and minimized the errors of several thermodynamic quantities.\(^{82}\) Zhang and Wang constructed a variant of the long-range-corrected Perdew–Burke–Ernzerhof functional based on the localized orbital locator (LC-\( \omega \)PBE\(^{LOL} \)) that was defined using the kinetic energy, and predicted the singlet–triplet gap (\( \Delta E_{\text{ST}} \)) of metal-free TADF emitters in comparable quality to OT-\( \omega \)PBE with a small mean absolute error (MAE = 0.09 eV).\(^{83}\) Corminboeuf and coworkers designed a ML-based \( \omega \)PBE functional that recovers the exact DD using the piecewise relationship between the average energy curvature and the electronic number, and reduced the error of the fundamental gap (\( E_g \)) of large hole-transporting molecular materials from 0.54 eV by one version of LC-\( \omega \)PBE (\( \omega = 0.400 a_0^{-1} \)) to 0.15 eV.\(^{84}\) Compared to direct ML predictions of excited state properties,\(^{96–100}\) these ML-RSH schemes maintain rigorous solutions of time-(in)dependent Kohn–Sham equations along with valid physical meanings, although they suffer from overfitting and undergeneralization problems due to the limited size and diversity of training sets.

Herein we proposed an “top-down” ML strategy because it optimizes the molecule-specific \( \omega \) in a RSH based on inexpensive molecular descriptors rather than Koopmans’ theorem or the energy v.s. electron number curvature. Our representative functional ML-\( \omega \)PBE borrows the formula of LC-\( \omega \)PBE\(^{64} \) but leaves the value of \( \omega \) to be determined using ML. We will show that ML-\( \omega \)PBE reaches a comparably strong predictive power to OT-\( \omega \)PBE but only consumes a fractional time after complete training. To ensure the high generalizability and
accuracy of our model, we were very cautious in deciding the database, molecular descriptors, and algorithms.

For our database we needed molecules that possess rich spectroscopic and photochemical properties but are notorious in theoretical studies due to SIE. We selected 3,926 such species from existing databases, including 1,941 solar cell materials from the Harvard Clean Energy Project (CEP), 101,102 904 pharmaceutically significant compounds from the DeepChem database, 103 431 fluorescence species from the ChemFluor database, 104 337 organic photovoltaic (OPV) molecules from the Harvard Organic Photovoltaic Dataset (HOPV15), 105 84 organic light-emitting diode (OLED) materials studied by Aspuru-Guzik and coworkers, 106 and 229 oligomers added by us in the present work. These compounds were randomly distributed into a training set of 1,970 and a test set of 1,956, and their structures were provided as simplified molecular-input line-entry system (SMILES) strings in the Supporting Information (SI). We generated the optimal value of $\omega$ for each molecule using OT-$\omega$PBE ($\omega_{\text{OT}}$). For the training set these $\omega_{\text{OT}}$’s were included in the training data along with molecular descriptors, while for the test set they were utilized as a reference to calibrate ML-$\omega$PBE. A wide distribution of these $\omega_{\text{OT}}$’s was shown to be between 0.120$a_0^{-1}$ and 0.330$a_0^{-1}$ for the entire database (Fig. 1), with statistics of $\bar{\omega}_{\text{OT}} = (0.206 \pm 0.029)a_0^{-1}$. Different databases presented distinct ranges of $\omega_{\text{OT}}$, but only 10 species lied at $\omega_{\text{OT}} > 0.300a_0^{-1}$. This result confirmed the diversity of our database but meanwhile invalidated the choice of $\omega = 0.300a_0^{-1}$ in the default LC-$\omega$PBE 64 to the majority of organic semiconducting molecules.

To describe important molecular features, we needed a molecular descriptor that captures unique structural and electronic properties. We herein introduced a “composite molecular descriptor” which implements contributions from several singular molecular descriptors like combined molecular fingerprints (CMF), 107–111 physical organic descriptors (POD), 108,109 and electronic structure properties (ESP). 112–114 Detailed information about the singular and composite molecular descriptors were provided in the SI. A precursor of our composite molecular descriptor was designed by Ju and coworkers and proved a success. 104 In a pre-
Figure 1: $\omega_{\text{OT}}$ evaluated using Koopmans’-theorem-based OT-$\omega$PBE for all 3,926 organic semiconducting molecules in the database, labeled by their sources. The default LC-$\omega$PBE with $\omega = 0.300a_0^{-1}$ (vertical dashed line) is a poor representation for the present dataset.

liminary model of ML-$\omega$PBE constructed using the extreme gradient boosting (XGBoost) algorithm, our composite molecular descriptor (CMF + POC + ESP) outperformed all singular or dual molecular descriptors in predicting $\omega_{\text{ML}}$ for the test set (Table S3), exhibiting the lowest mean absolute errors: MAE = 0.00573$a_0^{-1}$ for composite, $\geq 0.00621a_0^{-1}$ for singular and $\geq 0.00577a_0^{-1}$ for dual, with all errors defined as $\Delta = \omega_{\text{ML}} - \omega_{\text{OT}}$. This result endorsed our choice of the composite molecular descriptor.

To enhance the overall predictive power of ML-$\omega$PBE and future ML-RSH functionals, we desired an algorithm that supersedes simple regression models and artificial neural networks in terms of promoting the predictive power and mitigating the overfitting problem. Earlier methodological studies established the stacked ensemble machine learning (SEML) approach on top of multiple regression methods (base learners) using an overall analytical algorithm (meta learner), and proved them to meet our demands. For our present SEML protocol (Fig. 2), we selected eight base learners, including two versions of least absolute shrinkage and selection operator (LASSO_1 and LASSO_2), random forest (RF), gradient boosted regression trees (GBRT), XGBoost, light gradient boosting machine (Light-GBM), kernel ridge regression (KRR), and support vector machine (SVM).
The criteria of selecting base learners include (a) their accuracy and (b) their abilities to
differentiate between each other. Criterion (a) makes sure a base learner provides effective
information to the meta learner, and criterion (b) guarantees that the stacked ensemble
learning algorithm improves the predictive power over every single base learner. Therefore,
we chose tree-based algorithms, kernel-based algorithms, and linear algorithms. Each base
learner generated a quantitative relationship between the composite molecular descriptor
and the optimal value of $\omega$ though the out-of-fold (OOF) cross-validation.\textsuperscript{137} After that
the least angle regression (LARS)\textsuperscript{138} algorithm was used as the overall meta learner, which
analyzed all trained relationships, exploited advantages of all base learners, and optimized
the combined results.\textsuperscript{116-121} Given the rate-limiting step of SEML to be the low-level electron-
structure calculation (GFN$n$-xtb in Fig. 2), the meta learner does not add significant
computational costs to the entire procedure. In a performance analysis of ML-$\omega$PBE in the
accuracy of $\omega_{ML}$ for the test set (Fig. 3(a)(b) and Table S5), our SEML model is superior
to every single base learner by giving the lowest errors in $\omega_{ML}$ ($\text{MAE} = 0.00504a_0^{-1}$), the
strongest linear correlation between $\omega_{ML}$ and $\omega_{OT}$ ($R^2 = 0.930\textsuperscript{139}$), as well as the smallest
occurrence of large errors: only 13.2\% of molecules showed $|\Delta| > 0.010a_0^{-1}$ and 2.6\% pre-
sented $|\Delta| > 0.020a_0^{-1}$. Compared to OT-$\omega$PBE, our SEML-based ML-$\omega$PBE reduced the
average computational cost in determining $\omega$ for the test set by 2.66 orders of magnitude,
from 41,940 s to 92 s for the test set. Our analysis supported the earlier assertion that SEML
surpasses all individual base learners regarding applicability and predictability.

Following the discussions of the ML strategy and the accuracy of $\omega_{ML}$, we will validate
the potential of ML-$\omega$PBE in high-throughput materials and drug discovery by showing
that its generalizability and reliability is quantitatively comparable to OT-$\omega$PBE\textsuperscript{140} and is
more substantial than other popular functionals, including four non-tuned RSH function-
als: LC-$\omega$PBE with $\omega = 0.300a_0^{-1}$,\textsuperscript{64} CAM-B3LYP,\textsuperscript{47} M06-2X,\textsuperscript{141} and $\omega$B97X-D3\textsuperscript{54} and
four non-RSH functionals: HF,\textsuperscript{142} PBE,\textsuperscript{143} PBE0,\textsuperscript{40,41} and B3LYP.\textsuperscript{37-39} We will evaluate
their performances on multiple experimentally observables including $E_g$, the vertical absorp-
The model reads structural and electronic properties, translates them into composite molecular descriptors, and optimizes $\omega$ based on eight base learners and the meta learner of LARS. 

The majority of molecules in the present test set of optical properties, such as all AIE compounds and many OPV and TADF structures were not included in the above-mentioned test set of $\omega_{ML}$, so that they served as great benchmarks for the generalizability of our model. The statistics of errors, defined based on the signed differences between calculated and experimental values, are summarized in Figs. 4 and S2, as well as Tables S6 through S9.

$E_g$ is the most straightforward observable to assess because it is the difference between two important eigenvalues that are relevant to Koopmans' theorem, $\varepsilon_{HOMO}$ and $\varepsilon_{LUMO}$ (Eq. (2)), and is susceptible to the quality of the asymptotic density decay. As we expected (Fig. 4 and Table S7), non-RSH functionals that had been successful on small organic molecules were no longer accurate for medium and large organic semiconducting molecules (MAE $\geq$ 2.401...
Figure 3: (a) Comparison in the error of $\omega_{\text{ML}} (\Delta)$ optimized using eight base learners and the SEML algorithm. (b) Comparison between $\omega_{\text{OT}}$ (x-axis) and $\omega_{\text{ML}}$ (y-axis) with labels of $|\Delta| = 0.000$ (black line), $0.010a_0^{-1}$ (dark gray lines) and $0.020a_0^{-1}$ (light gray lines). Our SEML-based ML-$\omega$PBE presents the lowest MAE of $0.00504a_0^{-1}$ and the lowest RMSE of $0.00761a_0^{-1}$, as well as the smallest percentages for large errors (13.2% for $|\Delta| > 0.010a_0^{-1}$ and 2.6% for $|\Delta| > 0.020a_0^{-1}$). The average computational cost of ML-$\omega$PBE is 2.66 orders of magnitude smaller than OT-$\omega$PBE.
eV), while the performance of any RSH functional was more acceptable (MAE ≤ 1.209 eV) especially ML-ωPBE (MAE = 0.681 eV) and OT-ωPBE (MAE = 0.733 eV). The advantage was magnified for molecules with large spatial extents of π-conjugations like OPV (MAE = 0.441 eV, Fig. S3(a)) and PAH (MAE = 0.271 eV, Fig. S3(b)) whose energetics are keen to the choice of ω, but vanished with smaller π-bonds like BIO (MAE = 1.870 eV, Fig. S3(c)) without fitting to experimental values. The surprisingly better behavior of ML-ωPBE than OT-ωPBE was probably due to the bias from the incomplete test subset or the cancellation between systematic and random errors, but it did not damage our conclusion about the accuracy of ML-ωPBE.

Figure 4: MAEs of $E_g$, $E_{abs}$, $E_{fl}$, $E_{ph}$, and $\Delta E_{ST}$ (all in eV) were compared across various XC functionals for selected molecules in the test set. The lowest and second lowest MAEs were labeled with stars and circles respectively. For these organic semiconducting molecules, ML-ωPBE exhibits a quantitatively comparable accuracy to OT-ωPBE, especially for molecules with large π-conjugations and charge-separated excited states.

In spite of weaker relevance to Koopmans’ theorem, optical band gaps $E_{abs}$, $E_{fl}$, and $E_{ph}$ are more interesting benchmark quantities because they have more abundant and reliable experimental data from UV-vis spectra and their accurate predictions rely on electron–hole interactions in addition to asymptotic properties. They were all evaluated using linear-response time-dependent DFT (LR-TDDFT), with $E_{abs}$ at the ground state geometries
where the training occurred and $E_{\text{fl}}$ and $E_{\text{ph}}$ at corresponding excited state geometries. For these three optical band gaps, although hybrid non-RSH functionals like PBE0\textsuperscript{40,41} and B3LYP\textsuperscript{38,39} can outperform non-tuned RSHs due to error cancellation between density-overlocalized HF and density-overdelocalized (semi-)local functionals, they were generally less accurate than ML-\omega PBE and OT-\omega PBE. ML-\omega PBE illustrated the lowest MAE of 0.194 eV for $E_{\text{fl}}$ and the second lowest MAEs of 0.182 eV and 0.206 eV for $E_{\text{abs}}$ and $E_{\text{ph}}$, respectively. These errors were only marginally different from OT-\omega PBE which provided MAEs of 0.170 eV, 0.204 eV, and 0.215 eV for $E_{\text{abs}}$, $E_{\text{fl}}$, and $E_{\text{ph}}$, respectively (Fig. 4 and Tables S6 and S7). Interestingly, just like OT-\omega PBE, ML-\omega PBE minimized the systematic errors and made the overall errors more evenly distributed around zero, showing the second lowest MSEs of +0.106 eV, +0.109 eV, and +0.115 eV for $E_{\text{abs}}$, $E_{\text{fl}}$, and $E_{\text{ph}}$. This discovery confirmed that ML-\omega PBE trained over ground state configurations are robustly generalizable to low-lying excited state structures.

Among all test species, subsets of TADF exhibited MAEs of 0.139 eV, 0.194 eV, and 0.206 eV from ML-\omega PBE for $E_{\text{abs}}$, $E_{\text{fl}}$, and $E_{\text{ph}}$, respectively (Fig. S3(d)) and AIE showed MAE = 0.239 eV from ML-\omega PBE for $E_{\text{abs}}$, (Fig. S3(e)). They illustrated apparent preference towards ML-\omega PBE and OT-\omega PBE because their HOMO and LUMO are located at two different parts of the molecule or complex, which stretches the electronic density to the asymptotic limit, and is stable over slightly varied local electron–hole interaction. The result of the AIE subset was of particular interest because it indicated a great potential to extend our model from single molecules to molecular complexes or aggregates without extra treatment, even when no molecular complexes and aggregates were included in the training set. On the contrary, HOMO and LUMO for the OPV subset (with MAE = 0.147 eV from ML-\omega PBE for $E_{\text{abs}}$, Fig. S3(a)) are greatly overlapped in space and emphasizes the subtlety in the electron–hole interaction, allowing the error-canceling PBE0 to win. In addition to the argument about orbital configurations, the large discrepancy between ML-\omega PBE and OT-\omega PBE in OPV indicates a second possible but resolvable origin of the error: a poor
prediction of $\omega$ in ML-$\omega$PBE due to the lack of electron acceptors in the training set.

In addition to optical band gaps, $\Delta E_{ST}$ is an essential property for the TADF subset because it is explicitly correlated to the quantum yield of TADF$^{146,152}$ and is highly responsive to the extent of charge separation. $\Delta E_{ST}$ exhibited a different but understandable trend from $E_{fl}$ or $E_{ph}$ (Fig. 4 and Table S8): many functionals that had significantly overestimated $E_{fl}$ and $E_{ph}$ (like M06-2X) or underestimated them (like PBE) predicted equally accurate $\Delta E_{ST}$ to ML-$\omega$PBE and OT-$\omega$PBE due to the cancellation in systematic error between singlet and triplet states. Such accidentally good results with other functionals, again, do not compromise the advantage of ML-$\omega$PBE.

Finally, we benchmarked the performance of ML-$\omega$PBE in $\mu_{fl}$ and $\tau_{fl}$, which are key quantities to determine difficult-to-evaluate composite photodynamic properties like the fluorescence quantum yield$^{18,19}$ and the luminescence dissymmetry factor.$^{153,154}$ We selected only five RSH functionals in our comparison due to their successful descriptions of electronic density distributions (Fig. S4(a)(b) and Table S9), and evaluated both $\mu_{fl}$ and $\tau_{fl}$ using LR-TDDFT. Because of a successful prediction of excited state electronic density, ML-$\omega$PBE provided almost equally accurate prediction to OT-$\omega$PBE for both $\mu_{fl}$ (MAE = 0.830 debye versus 0.827 debye) and $\tau_{fl}$ (MAE = 1.89 ns versus 1.90 ns), and outperformed other RSH functionals. This result allowed us to safely assert that ML-$\omega$PBE best described photodynamic properties in the LR framework. However, because the LR approximation is not able to capture large electronic rearrangement during charge-transferred transitions, it might end up with a poor emissive geometry and might introduce a systematic error that consistently overestimates $\mu_{fl}$ while underestimating $\tau_{fl}$. This problem explained the sizable MSEs across all functionals and cannot be resolved without methodological treatment.$^{155–158}$

To conclude the present study, our ML-$\omega$PBE functional was established using a top-down SEML strategy based on eight base learners and the LARS meta learner, and it can accurately and efficiently reproduced the molecule-specific formula of Koopmans’ theorem-based OT-$\omega$PBE using computationally inexpensive structural and electronic properties, as
well as theoretically challenging optical properties using low-cost DFT and LR-TDDFT. To make sure of sufficient reliability and generalizability for our SEML method, we selected 3,926 structurally diverse and experimentally well-studied organic semiconducting molecules into our training and test sets, and extracted their composite molecular descriptors based on structural analyses and the GFNn-xtb semi-empirical calculations. After a two-step training through the base and meta learners, we constructed an explicit map between \( \omega_{ML} \) and the composite molecular descriptor, and obtained an outstanding agreement between \( \omega_{ML} \) and \( \omega_{OT}^{140} \) for the test set, with an MAE of 0.00504 \( a_0^{-1} \) and a shorter timescale than OT-\( \omega \)-PBE by 2.66 orders of magnitude.

To further evaluate the quality of ML-\( \omega \)-PBE, we compared its performance on a few experimental observable optical properties with OT-\( \omega \)-PBE, four non-tuned RSH functionals, and four non-RSH functionals. Taking into account all possible error cancellations, the fully trained ML-\( \omega \)-PBE reached a wonderful performance that was quantitatively analogous to OT-\( \omega \)-PBE in all optical properties, and surpassed every single non-RSH and non-tuned RSH functional. The fact that ML-\( \omega \)-PBE was trained based solely on ground state properties and monomeric molecules did not weaken its predictive power on excited state properties and molecular aggregates or complexes. Our result validated the practical usefulness of ML-\( \omega \)-PBE in reproducing and predicting optical properties for real-life materials and drugs.

In conclusion, the present study did not only propose an accurate and efficient way to determine selected parameters in the XC functional, but also expanded the horizon of ML applications in the framework of quantum chemistry. We expected adequately trained ML-\( \omega \)-PBE and future ML-RSH functionals to replace first-principles OT-RSHs in large-scale virtual materials and drug discovery to overcome the bottleneck in the computational cost.

**Supporting Information** The Supporting Information is available free of charge on the ACS Publications website at DOI: XXX/XXXXXXX. SMILES strings and \( \omega \) values for all 3,926 molecules in the training and test sets, brief proof of Koopmans’ theorem and asymptotic decay of electronic density, descriptions of details for general OT-\( \omega \)-PBE and ML-
ωPBE functionals, composite molecular descriptors, the SEML model, quantum chemical calculations, and summaries of statistics of errors of ML-ωPBE and other XC functionals in optical properties (PDF).

Acknowledgments Z. L. would like to thank the University of Massachusetts Amherst for offering start-up funds and the Massachusetts Green High Performance Computing Center (MGHPCC) for providing computational resources. Z. L., C.-W. J. and E. J. F. would also like to thank Ms. Lixue Cheng, Prof. Hui Guan, and Dr. Kun Yao, for scientific insights and discussions, as well as Mr. Junyi Gong for providing the benchmarked solvent-based data of the AIE subset.

References


(34) Dreuw, A.; Weisman, J. L.; Head-Gordon, M. Long-Range Charge-Transfer Excited


(77) Jacquemin, D.; Moore, B.; Planchat, A.; Adamo, C.; Autschbach, J. Performance of an


(93) Schmidt, J.; Benavides-Riveros, C. L.; Marques, M. A. L. Machine Learning the Physi-


(114) Pracht, P.; Caldeweyher, E.; Ehlert, S.; Grimme, S. A Robust Non-Self-Consistent Tight-Binding Quantum Chemistry Method for Large Molecules. **2019**.


(157) Zhu, Y.; Herbert, J. M. Self-Consistent Predictor/Corrector Algorithms for Stable and

Range-Separated Functional

\[ I = -e_{\text{HOMO}} \]

\[ \frac{1}{r} = \frac{1 - \text{erf}(\omega r)}{r_{12}} + \frac{\text{erf}(\omega r)}{r} \]

Optimal Tuning  Machine Learning

Four Orders of Magnitude Faster!