Accurate Electronic and Optical Properties of Organic Doublet Radicals Using Machine Learned Range-Separated Functionals

Cheng-Wei Ju,†,‡ Yili Shen,¶,§ Ethan J. French,†,∥ Jun Yi,† Hongshan Bi,† Aaron Tian,†,⊥ and Zhou Lin*,†

†Department of Chemistry, University of Massachusetts Amherst, MA 01003, United States
‡Pritzker School of Molecular Engineering, The University of Chicago, Chicago, IL, 60637, United States
¶Manning College of Information and Computer Sciences, University of Massachusetts Amherst, MA 01003, United States
§College of Software Engineering, Tongji University, Yangpu, Shanghai 200092, China
∥Department of Mathematics and Statistics, University of Massachusetts Amherst, MA 01003, United States
⊥Massachusetts Academy of Math and Science, Worcester, MA, 01605, United States

E-mail: zhoulin@umass.edu
Abstract

Luminescent organic semiconducting doublet-spin radicals are unique and emergent optical materials because their fluorescent quantum yields ($\Phi_\text{fl}$) are not compromised by spin-flipping intersystem crossing (ISC) into any dark high-spin states. The multi-configurational nature of radical electronic structures challenges computational studies in the framework of single-reference density functional theory (DFT) and introduces room for method improvement. In the present study, we extended our earlier development of a machine-learned range-separated hybrid functional, referred to as ML-$\omega$PBE, from closed-shell molecules to doublet-spin radicals, and assessed its performance for the original training set of 3,926 organic semiconducting molecules and an external test set of 64 organic semiconducting radicals from five categories. Interestingly, for this external test set, ML-$\omega$PBE reproduced the optimal value of $\omega$, the molecule-dependent range-separation parameter, from the first-principles OT-$\omega$PBE functional with a small mean absolute error (MAE) of 0.0197 $a_0^{-1}$ and with a significant save of computational cost by 2.46 orders of magnitude. This result demonstrated excellent generalizability and transferability of ML-$\omega$PBE among a variety of organic semiconducting species. To further assess the predictive power of ML-$\omega$PBE on organic semiconducting radicals, we also compared its performance on experimentally measurable absorption and fluorescence energies ($E_{\text{abs}}$’s and $E_{\text{fl}}$’s), evaluated using time-dependent DFT (TDDFT), with nine conventional functionals. ML-$\omega$PBE reproduced experimental $E_{\text{abs}}$’s and $E_{\text{fl}}$’s for most radicals in questions, with small MAEs of 0.222 and 0.121 eV, marginally worse from OT-$\omega$PBE. Our work not only illustrated a successful extension of stacked ensemble machine learning (SEML) framework from closed-shell molecules to open-shell doublet-spin radicals, but also opened the venue for the calculations of optical properties these using single-reference TDDFT.
ML-$\omega$PBE

Doublet Radicals

$D_0$

$D_1$

$E_{abs}$

$E_{em}$
Organic semiconducting doublet-spin radicals can stabilize their unpaired electrons through the delocalized $\pi$-conjugations and can exhibit non-conventional non-Aufbau configurations where the singly occupied molecular orbitals (SOMO) are lower-lying than the highest (doubly) occupied molecular orbitals (HOMO).\textsuperscript{1–10} The resulting long-lived configurations and compelling physicochemical properties make them promising functional materials for emergent scientific fields, such as molecular magnetism\textsuperscript{11} and charge conductivity.\textsuperscript{7} In the present study, we are interested in radicals that function through controllable optical properties. For example, in the photothermal therapy (PTT), the radical anion of the supramolecular complex of benzodithiophene-fused perylene diimide (BPDI) and cucurbit[7]uril (CB[7]) absorbs at the biologically transparent near-infrared (NIR) wavelength and dissipates the photon energy as heat.\textsuperscript{12–17} In an organic light-emitting diode (OLED) device, the doublet first excited state ($D_1$) and can potentially reach a $100\%$ fluorescent quantum yield ($\Phi_{fl}$) because it does not undergo any spin-flipping intersystem crossing (ISC) into a high-spin dark state.\textsuperscript{4,7,18–20}

Optical properties of these radicals depend on geometric and electronic configurations associated with doublet ground and excited states ($D_n$). However, highly accurate theoretical treatment for $D_n$ states are challenging due to their open-shell characters.\textsuperscript{21} Many multi-configurational approaches have been developed based on density functional theory (DFT) recently to account for this problem while maintaining reduced computational costs. Examples include multi-configuration pair DFT (MC-PDFT),\textsuperscript{22} spin-adapted time-dependent DFT (X-TDDFT),\textsuperscript{23–26} spin-flip TDDFT (SF-TDDFT),\textsuperscript{27–29} orbital optimization DFT (OODFT),\textsuperscript{30–33} and frozen–density-embedding TDDFT (FDE-TDDFT).\textsuperscript{34} Compared to single-reference DFT and TDDFT, these approaches proved more physically correct and reliable. However, their applications are limited to small and simple systems because of the difficulties in selecting appropriate active spaces and electronic configurations without prior knowledge of the systems, as well as the less friendly computational time scales ($\simeq N_{\text{act}} N_{\text{orb}}^4$).\textsuperscript{35–39}

On the other hand, although being significantly limited by the theoretical challenge of the theory and the rareness of reliable experimental benchmarks, single-reference DFT...
and TDDFT exhibit lower computational costs ($\simeq N_{\text{orb}}^3$) and black-box characters and can produce electronic and optical properties of radicals to the desired accuracy if careful development and calibration have been performed for exchange–correlation (XC) functionals in advance.\textsuperscript{23,24,40–53} As outstanding examples, Head-Gordon and coworkers performed systematic studies for excited state properties of polycyclic aromatic hydrocarbon (PAH) radicals and\textsuperscript{43,44,50,54–63} other small radicals\textsuperscript{64–69} using original TDDFT and its modification with Tamm–Dancoff approximation\textsuperscript{54} (TDDFT/TDA) along with common XC functionals like BLYP\textsuperscript{70,71} and B3LYP.\textsuperscript{70–72} They found that TDDFT and TDDFT/TDA were able to reproduce experimental excited states with errors smaller than 0.3 eV when the basis set was reasonably large regardless the inexact functionals and adiabatic approximations or the inability to treat double excitation characters,\textsuperscript{73,74} and TDDFT/TDA outperformed TDDFT in capturing correct states by overcoming some orbital instability problems.\textsuperscript{75} They also managed to assign the strongest absorptions of these radicals to involve their SOMOs. For another set of examples, Tureček and coworkers investigated bright excited states of peptide radicals and reproduced their UV-vis and IR spectra using TDDFT and various range-separated hybrid (RSH) XC functionals.\textsuperscript{76,77} They found that the compatibility between the functional and the molecular character was very important: LC-BLYP\textsuperscript{78} and M06-2X\textsuperscript{79} exhibited the best performance for excitations associated to the $C_\alpha$ radicals, while $\omega$B97X-D\textsuperscript{59,80} was the most accurate for excitations associated with the $\pi$-systems.\textsuperscript{48,52} They also found that a peptide radical typically exhibited a high energy SOMO and a low-energy HOMO, so that its excitation energies from HOMO to SOMO was significantly greater than those from SOMO to LUMO.\textsuperscript{81–104} Some other researchers, such as Joblin,\textsuperscript{45,105–110} Jacquemin,\textsuperscript{47,111–116} Grimme,\textsuperscript{117–121} Furche,\textsuperscript{122–124} and Allouche,\textsuperscript{125–130} performed similar analyses on doublet radicals to benchmark functionals and basis sets and to obtain physical insights and reaction mechanisms. All these computational have demonstrated the advantages of using global hybrid (GH)\textsuperscript{70–72,79,131–140} or range-separated hybrid (RSH)\textsuperscript{80,141–151} XC functionals for organic
Having been aware of the demand of a molecule-dependent RSH XC functional for organic semiconductors, as well as the recent advancement in artificial intelligence (AI), we introduced a stacked ensemble machine learning (SEML) algorithm\textsuperscript{152–157} and designed a new functional referred to as ML-ωPBE.\textsuperscript{151} In that study, we determined the molecule-dependent range-separation parameter (ω), as defined in the separation of the Coulomb operator,

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

using a composite molecular descriptor (CMD).\textsuperscript{158–163} In addition, we systematically assessed the predictive power of ML-ωPBE in ω and multiple electronic and optical properties by comparing with many conventional functionals.\textsuperscript{70–72,79,132,135,141,146,164,165} Among these functionals include first-principles OT-ωPBE, where ω is optimally tuned based on Koopmans’ theorem\textsuperscript{150,151,166–173} by minimizing the metrics of

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

We proved that well-trained ML-ωPBE was equally as capable as OT-ωPBE in terms of the accuracy of ω with a mean absolute error (MAE) of 2.5%, but significantly reduced the computational cost by 2.66 orders of magnitude. In addition, ML-ωPBE reproduced experimental electronic and optical properties of interest as accurately as OT-ωPBE and exceeded the performance of every other functional. It was worth noticed that the training and test sets in our SEML model presented no overlaps. Especially, the test set included some “external” molecules with no structural analogues present in the training set.\textsuperscript{150,174–176} Successful treatments of these external species indicated advantages of SEML and ML-ωPBE that were seldom observed in other ML models and XC functionals, a strong transferability, or a substantial domain adaption.
Figure 1: Architecture of the SEML algorithm for ML-ωPBE and potential challenge in domain adaption from closed-shell molecules to doublet-spin radicals.

In the present study, we extended the development of ML-ωPBE from close-shell organic semiconducting molecules to doublet-spin organic semiconducting radicals (Figure 1). In particular, we addressed a greater and more worthwhile question about domain adaption: are we able to reproduce the success of ML-ωPBE on radicals? Herein we performed a similar performance assessment for ML-ωPBE based on 64 new species, including 35 carbon-based radicals (C), 2 polyaromatic hydrocarbon-based radicals (PAH), 13 nitrogen-based radicals (N), 6 nitrogen-oxygen-based radicals (NO), and 8 aryl oxygen-based radicals (ArO). We placed all these radicals in the new test set, and provided their XYZ coordinates in the Supporting Information (SI) for optimized D_0 and D_1 states. We also combined the training and test sets from the original study, a total of 3,926 molecules, into the new training set. We hypothesized that the absence of radical species from the training set does not weaken the predictive power of ML-ωPBE in ω and electronic and optical properties. To describe the structural and electronic configurations for all these molecules and radicals, we constructed their CMDs using similar inexpensive properties from the previous study. We revisited information about the new training set and all CMDs in SI.

In the present study, we reapplied the “top-down” SEML algorithm which implements the idea that the stacked generalization of several regression models (or base learners) demonstrate more power than every single model. Here we chose a total of eight
successful base learners from the literature of computer science and statistics.\textsuperscript{228,229,229–233}

Each base learner generated a non-linear quantitative relationship between the CMDs of a molecule and its optimal value of $\omega$ ($\omega_{\text{ML}}$). All these relationships were later fed into and analyzed by a master regression model (meta learner)\textsuperscript{235} which produced the final prediction of $\omega$. We provided information about the base and meta learners in the SI.
Figure 2: Values of $\omega_{ML}$ for 64 doublet-spin radicals in comparison with $\omega_{OT}$, illustrated using (a) distribution and (b) scattered plots. Average values $\langle \omega_{OT} \rangle = 0.178 \ a_0^{-1}$ and $\langle \omega_{ML} \rangle = 0.191 \ a_0^{-1}$ and the default value of $\omega = 0.300 \ a_0^{-1}$ in LC-$\omega$PBE are labeled using the dotted lines. $\omega_{ML} = \omega_{OT}$, $\omega_{OT} \pm \Delta \omega_{OT}$, $\omega_{OT} \pm 2\Delta \omega_{OT}$ are represented as dashed lines.
In the present study, we calibrated the performance of ML-\(\omega\)PBE in a few different aspects. First, we assessed the accuracy of ML-\(\omega\)PBE in the prediction of \(\omega\) and confirmed its capacity in domain adaption by showing that the non-linear quantitative relationship obtained from the domain of closed-shell molecules can be extrapolated to that of doublet-spin radicals. In Figures 2(a) and (b), we compared the values of \(\omega\) obtained from OT-\(\omega\)PBE and ML-\(\omega\)PBE (\(\omega_{\text{OT}}\) and \(\omega_{\text{ML}}\)) for the test set. From Figure 2(a), we showed that both \(\omega_{\text{OT}}\) and \(\omega_{\text{ML}}\) ranged broadly from 0.120 to 0.320 \(a_0^{-1}\), indicating that it is appropriate to implement system-dependent values of \(\omega\)'s rather than selecting any single universal value. Similar to the training set,\(^{151}\) distributions of both \(\omega_{\text{OT}}\) and \(\omega_{\text{ML}}\) demonstrated that the typical LC-\(\omega\)PBE (\(\omega = 0.300 a_0^{-1}\)) functional\(^{146}\) failed to capture correct electronic properties of radicals in the test set, because only the species with the smallest size, such as phenoxy (ArO-55), ever needed an optimal \(\omega > 0.300 a_0^{-1}\). In addition, Figure 2(a) provided statistics of \(\omega_{\text{OT}} = (0.178 \pm 0.039) a_0^{-1}\) and \(\omega_{\text{ML}} = (0.191 \pm 0.034) a_0^{-1}\), respectively, demonstrating shifts to lower values in comparison with the training set, confirming the more diffuse and delocalized electronic structures of doublet-spin radicals in the test set. From Figure 2(b), we found that most values of \(\omega_{\text{ML}}\)'s were very close to \(\omega_{\text{OT}}\)'s, arriving at a small MAE of \(\Delta_{\text{ML}} = 0.0197 a_0^{-1}\) with very narrow distribution of deviations: among all 64 radicals, 33 of them exhibited absolute errors \(\Delta_{\text{ML}} \leq \bar{\Delta}_{\text{ML}}\) and 61 \(\Delta_{\text{ML}} \leq 2\bar{\Delta}_{\text{ML}}\). Compared to closed-shell molecules reported in the previous study,\(^{151}\) this value \(\bar{\Delta}_{\text{ML}}\) was more than three times as large. However, we can still claim the successful domain adaption of ML-\(\omega\)PBE because (1) this value was only 11.1% of \(\langle \omega_{\text{OT}} \rangle\) and 10.3% of \(\langle \omega_{\text{ML}} \rangle\), which turned out not to affect the predictive power of ML-\(\omega\)PBE in electronic and optical properties for these radicals, and (2) the current training set was comprised of molecules only but no radicals. In addition, ML-\(\omega\)PBE generated a comparable save of computational cost for radicals: the average computational time to optimize \(\omega\) (\(\langle t_\omega \rangle\)) was reduced from 63,442 seconds for OT-\(\omega\)PBE to 221 seconds for ML-\(\omega\)PBE, by 2.46 orders of magnitude.
Figure 3: (a) t-SNE results on the training set of molecules (colorful spheres), as well as the test set of radicals (black cubes) along with their closed-shell hydrogenated counterparts (red tetrahedrons). All results were generated using the CMD constructed with ECFP4 (Morgan) and PaDEL fingerprints. The $x$- and $y$-axes represent two components from principal component analysis (PCA). The color bar represents the value of $\omega_{OT}$ for the training set. (b) Comparison in the value of $\omega_{ML}$ between radicals in the test set ($x$-axis) and their closed-shell hydrogenated counterparts ($y$-axis).
Second, we rationalized the successful domain adaption of ML-ωPBE. In our first effort, we analyzed the chemical space spanned by the CMD and occupied by the training and test sets using the t-distributed stochastic neighbor embedding (t-SNE)\textsuperscript{238} method, as illustrated by Figures 3(a) and S2. To extract essential information and validate the necessity of using CMDs, we used simplified CMDs constructed using ECFP4 (Morgan)\textsuperscript{159,236} and PaDEL\textsuperscript{237} fingerprints in Figure 3 and the simple ECFP4 (Morgan) fingerprint in Figure S2. The t-SNE result demonstrated obviously that the chemical space occupied by the test set of radicals was widely distributed as long as an appropriate CMD was applied, but did not exceed the range spanned by the training set of molecules. Such a significant overlap deciphered the reason behind the successful domain adaption. Also, we noticed that the results were more clustered when described using the CMD from ECFP4 and PaDEL compared to the simple ECFP4 fingerprint, validating the stronger capacity of differentiating molecules by a CMD than a single descriptor. As a further validation, we compared the values of ω\textsubscript{ML} evaluated for all 64 doublet-spin radicals in the test set and their closed-shell counterparts with an additional hydrogen atom added to the radical site (Figure 3(b)). We found that two groups of ω\textsubscript{ML} values were very close to each other, with a tiny MAE of 0.00434 a\textsubscript{0}^{-1}. Such an extreme similarity proved that the structural and electronic configurations of radicals and their hydrogenated counterparts are only marginally differentiable from each other.
Figure 4: (a) Chemical structures of AIE-16, TADF-8, C-11, C-12, and N-34. (b) Energies and structures of frontier MOs of C-11 (\(\omega = 0.160 \, a_0^{-1}\)), C-12 (\(\omega = 0.184 \, a_0^{-1}\)), and N-34 (\(\omega = 0.185 \, a_0^{-1}\)) from ML-\(\omega\)PBE. (c) Relationships between the calculated values of \(E_{\text{abs}}\) (eV) associated with \(S_0 \rightarrow S_1\) (AIE-16 and TADF-8) or \(D_0 \rightarrow D_1\) (C-11, C-12, and N-34) transitions and the scanned values of \(\omega\) (\(a_0^{-1}\)), with experimental values provided. (c) Relationships between calculated total spins (\(\langle S^2 \rangle \)) associated with \(D_0\) and \(D_1\) states of C-11, C-12, and N-34 and scanned values of \(\omega\) (\(a_0^{-1}\)).
Third, before we systematically tested the accuracy of ML-\(\omega\)PBE on electronic and optical properties of doublet-spin radicals, we benchmarked the relationship between the electronic configurations evaluated using LC-\(\omega\)PBE with varying values of \(\omega\).\textsuperscript{239,240} We were motivated to take this short detour by the lack of assessment of RSH XC functionals on open-shell systems. As presented in Figure 4(a), we selected two representative molecules from the training set, 2,2',2"-(2-phenylethene-1,1,2-triyl)trinaphthalene (AIE-16) with a locally excited (LE) first singlet excited state (S\(_1\))\textsuperscript{151,174} and 2-phenyl-5-(4-(10-phenylphenazin-5(10H)-yl)phenyl)-1,3,4-oxadiazole (TADF-8) with a charge transfer (CT) S\(_1\) state.\textsuperscript{151,176} We also picked three representative radicals from the present test set, including carbon-based radicals (2',3',5',6'-tetrafluoro-N,N-diphenyl-4-amine-[1,1'-biphenyl])-(bis(perchlorophenyl))methyl (C-11) with a primary CT D\(_1\) state, tris(3,5-diisopropylphenyl)methyl (C-12) with a primary LE D\(_1\) state, and the nitrogen-based radical S-(2,4-dichlorophenyl)-N-(5'-phenyl-[1,1':3',1"-terphenyl]-2'-yl)thiohydroxylaminyl (N-34) with a partial CT D\(_1\) character. For all or some of these five species, we illustrated their frontier molecular orbital (MO) configurations (Figure 4(b) and S3–S5), absorption energies (\(E_{\text{abs}}\), Figure 4(c)), total spin configurations (\(\langle S^2 \rangle \)) associated with the D\(_0\) and D\(_1\) states (Figure 4(d)), and natural transition orbital (NTO) pairs (Table S8 and Figures 4(e) and S6–S15) associated with the transitions of S\(_0\) (singlet ground state) \(\rightarrow\) S\(_1\) (molecules) or D\(_0\) \(\rightarrow\) D\(_1\) (radicals). Some of these quantities were evaluated as functions of \(\omega\) for 0.050 \(a_0^{-1} \leq \omega \leq 0.400 a_0^{-1}\).

Figures S3–S5 provided orbital energies for radicals of C-11, C-12, and N-34 as functions of \(\omega\), and Figure 4(b) showed the quantitative ordering of their energies evaluated at the typical value of \(\omega = 0.200 a_0^{-1}\) along with their shapes. Here the comparison between \(\beta\) and \(\alpha\) orbitals for each radical indicated the change in the electronic structures before and after the introduction of the unpaired single electron. Among these three radicals, C-12 gave an identical order of energies for frontier \(\beta\) and \(\alpha\) MOs, except that the occupied \(\alpha\) SOMO was significantly lowered from the unoccupied \(\beta\) SOMO. Interestingly the \(\alpha\) SOMO–LUMO gap was always energetically similar to the \(\beta\) HOMO–SOMO gap, and the spatial overlap
within each pair was large, suggesting a mixed character in the lowest-lying LE-like electron excitation. The other two radicals, C-11 and N-34, demonstrated re-ordered energies at different extents. For C-11, the energy of the occupied $\alpha$ SOMO was reduced substantially so that it went below $\alpha$ HOMO, exhibiting a non-aufbau configuration after introducing the unpaired electron. For N-34, the nearly degenerate HOMO and HOMO−1 switch order before and after introducing the unpaired electron. For both radicals, the $\beta$ SOMO–LUMO gap was obviously smaller than the $\alpha$ HOMO–LUMO gap, making $\beta$ SOMO $\rightarrow$ LUMO the leading contributor of the first excited state. The spatial overlap within each pair validated their CT-like and partial-CT-like characters. The NTO pairs from Figure S6–S15 confirmed these analyses.

Figure 4(c) exhibited the bimodal relationship between $E_{\text{abs}}$ and $\omega$: for AIE-16, TADF-8, and N-34, $E_{\text{abs}}$ monotonically increased with $\omega$ as expected, because the raised overall portion of Hartree–Fock (HF) exchange tended to over-localize electrons and over-estimate $E_{\text{abs}}$. Their leading NTO pairs from Figure S6, S7, and S15 remained similar across the broad range of $\omega$, except that the fraction of the CT character monotonically decreased with an increasing $\omega$, and small contributions (amplitude < 0.20) from $\alpha/\beta$ HOMO−1 $\rightarrow$ LUMO and $\alpha/\beta$ HOMO−1 $\rightarrow$ LUMO+2 appeared for AIE-16 and TADF-8, respectively. On the contrary, C-11 and C-12 did not demonstrate monotonic trends. Instead, their $E_{\text{abs}}$ increased first with rising value of $\omega$, peaked at $\omega = 0.310$ and 0.290 $a_0^{-1}$, respectively, and decreased afterwards. In addition to the ever-increasing localization, the characters of the $D_0 \rightarrow D_1$ transitions were more complicated due to the introduction of transition mixings and character shifts between 0.200 $a_0^{-1}$ and 0.300 $a_0^{-1}$. For C-11, the dominant NTO pair bore the character of $\beta$ HOMO $\rightarrow$ SOMO, but its amplitude dropped from 1.00 to 0.65 when $\omega$ increased from 0.050 $a_0^{-1}$ to 0.400 $a_0^{-1}$, while a secondary NTO pair, with the character of $\alpha$ HOMO $\rightarrow$ LUMO+1, increased its amplitude from 0.00 to 0.56. Similarly, C-12 possessed two dominant NTO pairs, with the characters of SOMO $\rightarrow$ LUMO ($\alpha$) and HOMO−3 $\rightarrow$ SOMO ($\beta$), gradually changed their amplitudes from 0.74 and 0.70 to 0.67 and 0.74,
respectively.

The spin configurations for C-11, C-12, and N-34 in Figure 4(d) supplied more some rationalization for the mixing of NTOs. Although we found a breaking of spin symmetry in all three radicals with an increasing fraction of HF exchange, photoexcitations played distinct roles on them. For N-34, the values of $\langle S^2 \rangle$ were negligibly different between $D_0$ and $D_1$ and were never far away from the expected value of 0.75 for a pure doublet radical. For C-11 and C-12, instead, the values of $\langle S^2 \rangle$ for $D_0$ were close to the expected value of 0.75, while those for $D_1$ reached as high as 1.80 for C-11 and 1.50 for C-12. The significant but gradual breakdown of the spin symmetry at the excited states of C-11 and C-12 agreed with the ever-increasing mixing character of NTO pairs in them and explained bimodal configurations for $E_{\text{abs}}$. The situations were more serious in C-11 because its orbitals were more delocalized and its transitions were more charge transferred.

All our discussions herein revealed an essential reason for optimizing $\omega$ for radicals, the difficulty and instability embedded in RSH functionals when being applied to open-shell systems. In particular, the excited-state electronic structures of doublet-spin radicals in terms of transition characters and spin symmetries were extremely sensitive to the choice of $\omega$, especially when they exhibited more delocalized or CT characters.

![Figure 5: MAEs (eV) of $E_{\text{abs}}$ and $E_{\text{fl}}$ were evaluated using ML-$\omega$PBE and TDDFT/6-311G(d) for selected radicals in the test subset and compared with eight conventional XC functionals.](image)

Finally, we commenced the benchmark study of ML-$\omega$PBE by examining its predictive power for $E_{\text{abs}}$ and the fluorescence energy ($E_{\text{fl}}$) for doublet-spin radicals in the test set. For $E_{\text{abs}}$ we constructed a test subset of 48 radicals with reliable experimental values to
compare, including 19 from C, 2 from PAH, 13 from N, 6 from NO, and 8 from ArO. Also for \( E \) we built up a test subset of 16 C radicals with trustable experimental measurements. We calculated \( E_{\text{abs}} \) and \( E_{\text{fl}} \) for corresponding radicals using the original TDDFT scheme and TDDFT/TDA, as well as different basis sets. We compared the accuracy of ML-\( \omega \)PBE in terms of MAEs and/or mean signed errors (MSEs) with eight existing XC functionals, including OT-\( \omega \)PBE,\textsuperscript{150,151} LC-\( \omega \)PBE with \( \omega = 0.200 \, \text{a}_0^{-1} \) and \( 0.300 \, \text{a}_0^{-1} \),\textsuperscript{146} CAM-B3LYP,\textsuperscript{165} \( \omega \)B97X-D3,\textsuperscript{141} M06-2X,\textsuperscript{79} PBE,\textsuperscript{164} PBE0,\textsuperscript{132,135} and B3LYP,\textsuperscript{70–72} and provided our statistics in Figure 5 and Tables S1–S7 in the SI. Based on these results, we drew a few important conclusions about the strong predictive power of ML-\( \omega \)PBE. To begin with, we confirmed the extremely high sensitivity of the accuracy of \( E_{\text{abs}} \) and \( E_{\text{fl}} \) to the value of \( \omega \), especially for radicals with CT-like \( D_1 \) states like C-11, because an optimal \( \omega \) provided an essential balance between over-delocalizing PBE and over-localizing HF. As we had expected, the standard LC-\( \omega \)PBE with \( \omega = 0.300 \, \text{a}_0^{-1} \) showed a poor performance regardless of the choice of the TDDFT variant and the basis set, because \( \omega = 0.300 \, \text{a}_0^{-1} \) was far from \( \langle \omega_{\text{OT}} \rangle = 0.178 \, \text{a}_0^{-1} \) and \( \langle \omega_{\text{ML}} \rangle = 0.191 \, \text{a}_0^{-1} \) obtained from all of doublet-spin radicals in question. On the other hand, if we reduced \( \omega \) to a value closer to \( \langle \omega_{\text{OT}} \rangle \) and \( \langle \omega_{\text{ML}} \rangle \), like \( \omega = 0.200 \, \text{a}_0^{-1} \), LC-\( \omega \)PBE demonstrated a significantly improved performance, but was not able to reach consistently comparable MAEs and MSEs with ML-\( \omega \)PBE and OT-\( \omega \)PBE without optimizing the values of \( \omega \).

This was particularly true for large C radicals for which the optimal values of \( \omega \) were substantially lower than 0.200 \( \text{a}_0^{-1} \), such as 4’-(9H-carbazol-9-yl)-2,3,5,6-tetrachloro-[1,1’-biphenyl]- (bis(perchlorophenyl)methyl (C-9) (\( \omega_{\text{ML}} = 0.162 \, \text{a}_0^{-1} \)), C-11 (\( \omega_{\text{ML}} = 0.160 \, \text{a}_0^{-1} \)), and tris(4- (9-butyl-9H-carbazol-3-yl)-2,3,5,6-tetrachlorophenyl)methyl (C-19) (\( \omega_{\text{ML}} = 0.137 \, \text{a}_0^{-1} \)). The sensitivity mentioned in this paragraph, once again, re-validated the necessity to apply a system-dependent value of \( \omega \) for organic semiconducting radicals.

Next, we showed that well-trained ML-\( \omega \)PBE outperformed conventional functionals and accurately reproduced experimental values of photophysical properties. We obtained a strong performance on \( E_{\text{abs}} \) for test subsets of C, PAH, N, and NO radicals, with an overall MAE of
0.222 eV and an overall MSE of +0.126 eV using TDDFT/6-311G(d). For these four subsets, the predictive power of ML-ωPBE was only marginally different to that of OT-ωPBE and exceeded all other XC functionals. A similar behavior was demonstrated for $E_n$. These successful behaviors were significantly attributed to the excellent agreement between $\omega_{ML}$ and $\omega_{OT}$ (Figure 2(b)), as well as the delocalized and/or CT characters, for most organic semiconducting radicals in question. Further, they implied that our CMDs and the SEML algorithm can precisely represent the structural and electronic configurations of these radicals and can reliably construct their relationships with optimal values of $\omega_{ML}$, further confirming their potentials to adapt themselves from one domain to the other with distinct properties. In order to illustrate our analysis, we compared the characters of frontier MOs for $\alpha$ electrons of C-9 generated by ML-ωPBE and every other XC functional in question, as well as the ab initio CASCI approach (Figure 6). We found that CASCI predicts an energy order of $HOMO-1 < HOMO < SOMO$ for $\alpha$ orbitals. ML-ωPBE ($\omega = 0.162 \, a_0^{-1}$), OT-ωPBE ($\omega = 0.173 \, a_0^{-1}$), and one C-ωPBE ($\omega = 0.200 \, a_0^{-1}$) slightly switching the order by giving $HOMO-1 < SOMO < HOMO$. Any other XC functional, on the other hand, significantly switched the order by giving $SOMO < HOMO-1 < HOMO$.

Among the non-RSH XC functionals in comparison, the global hybrid functional PBE0 with 75% PBE and 25% HF, appeared to be an exception, because it occasionally gave smaller MAEs and MSEs than ML-ωPBE and OT-ωPBE. This behavior was highly likely due to the error cancellation between the partitions of density-overdelocalizing PBE and density-overlocalizing HF, or between the energies of $D_0$ and $D_1$ states involved in the absorption and fluorescence transitions (Table S1–S7). However, combined with the significantly reordered frontier MOs reported by C-9 (Figures 6), we concluded that a great energy agreement does not necessary equal to a great description of the electronic structures. This observation also agreed with that made for closed-shell molecules in the previous study.

In addition, among all radicals in the test set, the ArO family consistently acquired poor reproduction of $E_{abs}$ regardless the XC functional being used, including ML-ωPBE and
OT-ωPBE. All of the functionals underestimated $E_{\text{abs}}$ by more than 1 eV, and the “best” behavior was given by ωB97X-D3, with an signed error of $-0.993$ eV using TDDFT/6-311G(d). We would show that this huge error originated from the incorrect or unstable electronic structures obtained from single-reference, DFT-based approach even on the ground state. To prove our hypothesis, we compared the values of $E_{\text{abs}}$ and orders of four frontier MOs (HOMO−1, HOMO, SOMO, and LUMO) of phenoxy (ArO-56), the smallest PhO radical, calculated by ML-ωPBE, all conventional functionals in question, and CASSCF and CASCI calculations (Figure S9). Based on our analysis, the errors associated with $E_{\text{abs}}$ obtained from ML-ωPBE and other conventional functionals were significantly greater than the benchmark CASSCF (10,7) calculations ($-0.145$ eV) due to the substantially reordered frontier MOs. For example, the occupied α SOMO (# 25α) predicted by CASSCF (10,7) exhibited a $\pi_z$ bond localized on the oxygen (O) atom and the carbon (C) atom next to it. However, ML-ωPBE ($\omega = 0.190 a_0^{-1}$) and OT-ωPBE ($\omega = 0.178 a_0^{-1}$) failed to reproduce a similar MO. Instead, its delocalized MOs HOMO−6 (# 18α) and HOMO−2 (# 22α) added exactly to this configuration. For another example, HOMO−7 (# 17α and # 17β) predicted by CASSCF (10,7) raised its relative position to HOMO (# 24α and # 24β) from ML-ωPBE and OT-ωPBE. All of the rest XC functionals exhibited similar behavior. These
results indicated fundamental problems in single-reference DFT, even with the most careful calibration, can produce serious errors to highly-localized organic semiconducting radicals.

In the end, we compared the performance of ML-ωPBE across different variants of the TDDFT approach and the basis set. We found that the basis sets that include the diffuse functions, such as 6-311G+(d), did not improve the accuracy of ML-ωPBE and any other XC functionals, because the orbitals associated with these radicals are not very delocalized. In Figures S6–S8, we compared HOMO−1, HOMO, and SOMO for C-11, C-12, and N-34 and their corresponding hydrogenated species, C-11-H, C-12-H, and N-34-H. The result indicated that the high-energy electrons in these radicals, including the unpaired ones, presented similar spatial occupation to their hydrogenated counterparts. In some situations they are even less delocalized. Also, we found that the inclusion of TDA slightly compromised the accuracy of both $E_{\text{abs}}$ and $E_{\text{fl}}$, indicating that these doublet radicals are less likely to suffer from the instability problems to some closed-shell molecules and it is necessary to include de-excitation and coupling matrices in the working eigenvalue equations of linear response.\(^{54,241–247}\)

In conclusion, we presented a follow-up assessment study for ML-ωPBE functional\(^{151}\) that was self-developed based on the top-down SEML strategy\(^{152–157}\) and expanded its application domain in the framework of DFT. Here we successfully demonstrated the capacity of ML-ωPBE in domain adaption from closed-shell singlet molecules\(^{222–227}\) in the training set to open-shell doublet radicals\(^{19,63,177–221}\) in the test set with distinct electronic structures. Even with no radicals present in the training set, ML-ωPBE reproduced the system-specific optimal values of $\omega$ generated by the Koopmans’ theorem-based OT-ωPBE with a MAE of 0.0197 $a_0^{-1}$ for all these radicals, but reduced the average computational cost by 2.46 orders of magnitude. Due to accurate captures of electronic structures, ML-ωPBE demonstrated an analogous top predictive power to OT-ωPBE in terms of two experimentally observable optical properties, $E_{\text{abs}}$ and $E_{\text{fl}}$, and outperformed every other XC functionals,\(^{70–72,79,132,135,141,146,164,165}\) without obvious error cancellations. The only exception was
the ArO family for which the single-reference DFT methods fail in general. In summary, we validated the practical value of ML-\(\omega\)PBE in deciphering and predicting optical properties for luminescent doublet-spin organic semiconducting radicals with potential applications in PTT and OLED and facilitated its application in large-scale computationally aided materials discovery.

Acknowledgement

The authors would like to acknowledge the UMass/URI Unity Cluster and MIT SuperCloud\textsuperscript{248} for providing high-performance computing (HPC) resources. Z.L. would like to acknowledge UMass Amherst for providing the start-up funds and the NSF-UMass ADVANCE Collaborative Research Seed Grant to accomplish this study. Also, we would like to thank helpful discussion from Prof. Yuezhi Mao, Prof. Hui Guan, Dr. Kun Yao, Dr. Lixue (Sherry) Cheng, and Mr. Junjie Yang.

Supporting Information Available

The Supporting Information is available free of charge at XXXXXXXXXX.

- Details of quantum chemical calculations; brief revisit of the SEML model; similarity and difference in chemical space between singlet molecules and doublet radicals; error statistics of ML-\(\omega\)PBE and other XC functionals in optical properties; and configurations of frontier MOs and NTOs (PDF).
- Optimized \(D_0\) and \(D_1\) geometries for all 64 radicals in the external test set (ZIP).
- SMILES strings and \(\omega\) values for all 64 radicals in the external test set (XLSX).
References


(8) Sun, Y.; Moe, M. M.; Liu, J. Mass Spectrometry and Computational Study of Collision-Induced Dissociation of 9-Methylguanine–1-Methylcytosine Base-Pair Radical Cation: Intra-Base-Pair Proton Transfer and Hydrogen Transfer, Non-Statistical


(37) Evangelista, F. A. Perspective: Multireference Coupled Cluster Theories of Dynamical Electron correlation. 

(38) Sun, Q.; Yang, J.; Chan, G. K.-L. A General Second Order Complete Active Space Self-Consistent-Field Solver for Large-Scale Systems. 


(41) Tokmachev, A. M.; Boggio-Pasqua, M.; Mendive-Tapia, D.; Bearpark, M. J.; Robb, M. A. Fluorescence of the Perylene Radical Cation and an Inaccessible $D_0/D_1$ Conical Intersection: An MMVB, RASSCF, and TD-DFT Computational Study. 
(42) Chillier, X. D. F.; Stone, B. M.; Joblin, C.; Salama, F.; Allamandola, L. J. \(D_1(^2B_{3g}) \rightarrow D_0(^2A_u)\) Fluorescence from the Matrix-Isolated Perylene Cation Following Laser Excitation into the \(D_5(^2B_{3g})\) and \(D_2(^2B_{3g})\) Electronic States. *J. Chem. Phys.* **2002**, *116*, 5725–5730.


(49) Yu, X.; Hou, H.; Wang, B. Atmospheric Chemistry of Perfluoro-3-Methyl-2-Butanone


(63) Small, D.; Zaitsev, V.; Jung, Y.; Rosokha, S. V.; Head-Gordon, M.; Kochi, J. K. Intermolecular π-to-π Bonding between Stacked Aromatic Dyads. Experimental and Theoretical Binding Energies and Near-IR Optical Transitions for Phenalenyl Rad-


(69) Peverati, R.; Bera, P. P.; Lee, T. J.; Head-Gordon, M. Insights into Hydrocarbon Chain and Aromatic Ring in the Interstellar Medium: Computational Study of the Isomers of $\text{C}_4\text{H}_3^+$, $\text{C}_6\text{H}_3^+$ and $\text{C}_6\text{H}_5^+$ and Their Formation Pathways. *Astrophys. J.* **2016**, *830*, 128.


(95) Chung, T. W.; Moss, C. L.; Zimnicka, M.; Johnson, R. S.; Moritz, R. L.; Tureček, F.


(102) Shaffer, C. J.; Pepin, R.; Tureček, F. Combining UV Photodissociation Action Spec-


(126) Antoine, R.; Joly, L.; Allouche, A.-R.; Broyer, M.; Lemoine, J.; Dugourd, P. Elec-


Grimme, S.; Bannwarth, C.; Shushkov, P. A Robust and Accurate Tight-Binding Quantum Chemical Method for Structures, Vibrational Frequencies, and Noncovalent


(169) Stein, T.; Eisenberg, H.; Kronik, L.; Baer, R. Fundamental Gaps in Finite Systems


(192) Neckers, D.; Rajadurai, S.; Valdes-Aguilera, O.; Zakrzewski, A.; Linden, S. Photo-
physics and Photochemistry of P-Benzoylphenyldiphenylmethyl in Solution. *Tetrahe-

(193) Ruberu, S. R.; Fox, M. A. Photochemical Behavior of Stable Free Radicals: The
149.

(194) Peng, Q.; Obolda, A.; Zhang, M.; Li, F. Organic Light-Emitting Diodes Using a
**2015**, *54*, 7091–7095.

(195) Cui, Z.; Ye, S.; Wang, L.; Guo, H.; Obolda, A.; Dong, S.; Chen, Y.; Ai, X.; Ab-
durahman, A.; Zhang, M. et al. Radical-Based Organic Light-Emitting Diodes with
6648.

(196) Jin, Q.; Chen, S.; Sang, Y.; Guo, H.; Dong, S.; Han, J.; Chen, W.; Yang, X.; Li, F.;
Duan, P. Circularly Polarized Luminescence of Achiral Open-Shell $\pi$-Radicals. *Chem.

(197) Hattori, Y.; Kusamoto, T.; Nishihara, H. Highly Photostable Luminescent Open-Shell
(3,5-Dihalo-4-Pyridyl)bis(2,4,6-Trichlorophenyl)Methyl Radicals: Significant Effects
of Halogen Atoms on Their Photophysical and Photochemical Properties. *RSC Adv.*

(198) Uchida, K.; Mou, Z.; Kertesz, M.; Kubo, T. Fluxional $\sigma$-Bonds of the 2,5,8-
Trimethylphenalenyl Dimer: Direct Observation of the Sixfold $\sigma$-Bond Shift via a

(199) Patrascu, B.; Lete, C.; Popescu, C.; Matache, M.; Paun, A.; Madalan, A.; Ionitaa, P.


