Electronic Supplementary Information

Computational and Data Driven Molecular Material Design Assisted by Low Scaling Quantum Mechanics Calculations and Machine Learning

Wei Li*[a], Haibo Ma*[a,b], Shuhua Li[a], and Jing Ma*[a,b]

[a.] *Key Laboratory of Mesoscopic Chemistry of Ministry of Education, Institute of Theoretical and Computational Chemistry, School of Chemistry and Chemical Engineering, Nanjing University, Nanjing 210023, China*

[b.] *Jiangsu Key Laboratory of Advanced Organic Materials, Jiangsu Key Laboratory of Vehicle Emissions Control, Nanjing University, Nanjing 210023, China*

[c.] ** Email: majing@nju.edu.cn (J.M.); wli@nju.edu.cn (W.L.); haibo@nju.edu.cn (H.M.);*

List of Contents

Table S1: Selected works of using low-scaling QM methods to study molecular materials.

Low scaling QM method	Systems/Properties	References
Density matrix perturbation theory (DMPT)	Electric polarizability	Weber, V.; Niklasson, A. M. N.; Challacombe, M. Ab Initio Linear Scaling Response Theory: Electric Polarizability by Perturbed Projection. *Phys. Rev. Lett.* **2004**, *92* (19), 193002.
Density-functional perturbation theory (DFPT)	Electric polarizability	Shang, H.; Liang, W.; Zhang, Y.; Yang, J. Efficient Parallel Linear Scaling Method to Get the Response Density Matrix in All-Electron Real-Space Density-Functional Perturbation Theory. *Comput. Phys. Commun.* **2021**, *258*, 107613.
DM-based TDSCF	Hyperpolarizabilities	Kussmann, J.; Ochsenfeld, C. A Density Matrix-Based Method for the Linear-Scaling Calculation of Dynamic Second- and Third-Order Properties at the Hartree-Fock and Kohn-Sham Density Functional Theory Levels. *J. Chem. Phys.* **2007**, *127* (20), 204103.
Divide-and-Conquer (DAC)	Hyperpolarizabilities	Kobayashi, M.; Touma, T.; Nakai, H. Dynamic Hyperpolarizability Calculations of Large Systems: The Linear-Scaling Divide-and-Conquer Approach. *J. Chem. Phys.* **2012**, *136* (8), 084108.
CIM	Geometry optimization	Ni, Z.; Wang, Y.; Li, W.; Pulay, P.; Li, S. Analytical Energy Gradients for the Cluster-in-Molecule MP2 Method and Its Application to Geometry Optimizations of Large Systems. *J. Chem. Theory Comput.* **2019**, *15* (6), 3623–3634.
Divide-and-Conquer (DAC)	Geometry optimization	Kobayashi, M.; Kunisada, T.; Akama, T.; Sakura, D.; Nakai, H. Reconsidering an Analytical Gradient Expression within a Divide-and-Conquer Self-Consistent Field Approach: Exact Formula and Its Approximate Treatment. *J. Chem. Phys.* **2011**, *134* (3), 34105.
GEBF	Geometry optimization & Vibrational spectra	Hua, W.; Fang, T.; Li, W.; Yu, J. G.; Li, S. Geometry Optimizations and Vibrational Spectra of Large Molecules from a Generalized Energy-Based Fragmentation Approach. *J. Phys. Chem. A* **2008**, *112* (43), 10864–10872.
DCMB	Geometry optimization & Vibrational spectra	Wu, A.; Xu, X. DCMB That Combines Divide-and-Conquer and Mixed-Basis Set Methods for Accurate Geometry Optimizations, Total Energies, and Vibrational Frequencies of Large Molecules. *J. Comput. Chem.* **2012**, *33* (16), 1421–1432.
EE-GMFCC	Geometry optimization & Vibrational spectra	Liu, J.; Zhang, J. Z. H.; He, X. Fragment Quantum Chemical Approach to Geometry Optimization and Vibrational Spectrum Calculation of Proteins. *Phys. Chem. Chem. Phys.* **2016**, *18* (3), 1864–1875.
Hybrid many-body interaction (HMBI)	Lattice energies	Beran, G. J. O.; Nanda, K. Predicting Organic Crystal Lattice Energies with Chemical Accuracy. *J. Phys. Chem. Lett.* **2010**, *1* (24), 3480–3487.
GEBF	Lattice energies and crystal structures	Fang, T.; Li, W.; Gu, F.; Li, S. Accurate Prediction of Lattice Energies and Structures of Molecular Crystals with Molecular Quantum Chemistry Methods. *J. Chem. Theory Comput.* **2015**, *11* (1), 91–98.
SMF	Lattice dynamics	Collins, M. A. Ab Initio Lattice Dynamics of Nonconducting Crystals by Systematic Fragmentation. *J. Chem. Phys.* **2011**, *134* (16), 164110.
GEBF	IR/Raman spectra	Fang, T.; Jia, J.; Li, S. Vibrational Spectra of Molecular Crystals with the Generalized Energy-Based Fragmentation Approach. *J. Phys. Chem. A* **2016**, *120* (17), 2700–2711.
EE-GMFCC	IR/Raman spectra	Liu, J.; He, X. Fragment-Based Quantum Mechanical Approach to Biomolecules, Molecular Clusters, Molecular Crystals and Liquids. *Phys. Chem. Chem. Phys.* **2020**, *22* (22), 12341–12367.
MIM	VCD spectra	Jose, K. V. J.; Beckett, D.; Raghavachari, K. Vibrational Circular Dichroism Spectra for Large Molecules through Molecules-in-Molecules Fragment-Based Approach. *J. Chem. Theory Comput.* **2015**, *11* (9), 4238–4247.
GEBF	AIMD	Li, S.; Li, W.; Ma, J. Generalized Energy-Based Fragmentation Approach and Its Applications to Macromolecules and Molecular Aggregates. *Acc. Chem. Res.* **2014**, *47* (9), 2712–2720.
GMBE	AIMD	Liu, J.; Rana, B.; Liu, K.-Y.; Herbert, J. M. Variational Formulation of the Generalized Many-Body Expansion with Self-Consistent Charge Embedding: Simple and Correct Analytic Energy Gradient for Fragment-Based Ab Initio Molecular Dynamics. *J. Phys. Chem. Lett.* **2019**, *10* (14), 3877–3886.
CFM	NMR	Tan, H.-J.; Bettens, R. P. A. Ab Initio NMR Chemical-Shift Calculations Based on the Combined Fragmentation Method. *Phys. Chem. Chem. Phys.* **2013**, *15* (20), 7541.

GEBF	NMR	Zhao, D.; Song, R.; Li, W.; Ma, J.; Dong, H.; Li, S. Accurate Prediction of NMR Chemical Shifts in Macromolecular and Condensed-Phase Systems with the Generalized Energy-Based Fragmentation Method. *J. Chem. Theory Comput.* **2017**, *13* (11), 5231–5239.
MIM	NMR	Jose, K. V. J.; Raghavachari, K. Fragment-Based Approach for the Evaluation of NMR Chemical Shifts for Large Biomolecules Incorporating the Effects of the Solvent Environment. *J. Chem. Theory Comput.* **2017**, *13* (3), 1147–1158.
GEBF	NMR	D. Zhao, X. Shen, Z. Cheng, W. Li, H. Dong, S. Li, *J. Chem. Theory Comput.,* 2020, **16**, 2995-3005.
density matrix trace correcting (TC) purification	Boron nitrogen nanotubes	Luo, Z.; Qin, X.; Wan, L.; Hu, W.; Yang, J. Parallel Implementation of Large-Scale Linear Scaling Density Functional Theory Calculations With Numerical Atomic Orbitals in HONPAS. *Front. Chem.* **2020**, *8*, 932.
Pole expansion and selected inversion (PEXSI)	Graphene nanoflakes	Hu, W.; Lin, L.; Yang, C.; Yang, J. Electronic Structure and Aromaticity of Large-Scale Hexagonal Graphene Nanoflakes. *J. Chem. Phys.* **2014**, *141* (21), 214704.
GEBF	Supermolecular coordination complex	Yuan, D.; Li, Y.; Li, W.; Li, S. Structures and Properties of Large Supramolecular Coordination Complexes Predicted with the Generalized Energy-Based Fragmentation Method. *Phys. Chem. Chem. Phys.* **2018**, *20* (45), 28894–28902.
MIM	Supermolecular complex	Debnath, S.; Sengupta, A.; Jose, K. V. J.; Raghavachari, K. Fragment-Based Approaches for Supramolecular Interaction Energies: Applications to Foldamers and Their Complexes with Anions. *J. Chem. Theory Comput.* **2018**, *14* (12), 6226–6239.
GEBF	Supermolecular complex	Li, W.; Duan, M.; Liao, K.; Hong, B.; Ni, Z.; Ma, J.; Li, S. Improved Generalized Energy-Based Fragmentation Approach and Its Applications to the Binding Energies of Supramolecular Complexes. *Electron. Struct.* **2019**, *1* (4), 044003.
CIM	Supermolecular complex	Chen, Y.; Fu, L.; Sun, B.; Qian, C.; Wang, R.; Jiang, J.; Lin, C.; Ma, J.; Wang, L. Competitive Selection of Conformation Chirality of Water-Soluble Pillar[5]Arene Induced by Amino Acid Derivatives. *Org. Lett.* **2020**, *22* (6), 2266–2270.
Fermi operator expansion (FOE)	Organic molecules	Liang, W.; Saravanan, C.; Shao, Y.; Baer, R.; Bell, A. T.; Head-Gordon, M. Improved Fermi Operator Expansion Methods for Fast Electronic Structure Calculations. *J. Chem. Phys.* **2003**, *119* (8), 4117–4125.
Divide-and-conquer	Conjugated polymers	Kobayashi, M.; Nakai, H. How Does It Become Possible to Treat Delocalized and/or Open-Shell Systems in Fragmentation-Based Linear-Scaling Electronic Structure Calculations? The Case of the Divide-and-Conquer Method. *Phys. Chem. Chem. Phys.* **2012**, *14* (21), 7629.
CIM	zinc-organic batteries	Wang, Y.; Wang, C.; Ni, Z.; Gu, Y.; Wang, B.; Guo, Z.; Wang, Z.; Bin, D.; Ma, J.; Wang, Y. Binding Zinc Ions by Carboxyl Groups from Adjacent Molecules toward Long-Life Aqueous Zinc–Organic Batteries. *Adv. Mater.* **2020**, *32* (16), 2000338.
CIM		
MIM	Silicon cluster	Mayhall, N. J.; Raghavachari, K. Molecules-in-Molecules: An Extrapolated Fragment-Based Approach for Accurate Calculations on Large Molecules and Materials. *J. Chem. Theory Comput.* **2011**, *7* (5), 1336–1343.
Incremental method	Crystalline Silicon	Stoll, H. The Correlation Energy of Crystalline Silicon. *Chem. Phys. Lett.* **1992**, *191* (6), 548–552.
Local MP2	Nonconducting Crystals	Pisani, C.; Busso, M.; Capecchi, G.; Casassa, S.; Dovesi, R.; Maschio, L.; Zicovich-Wilson, C.; Schütz, M. Local-MP2 Electron Correlation Method for Nonconducting Crystals. *J. Chem. Phys.* **2005**, *122* (9), 094113.
DEC	Molecular crystals	Rebolini, E.; Baardsen, G.; Hansen, A. S.; Leikanger, K. R.; Pedersen, T. B. Divide–Expand–Consolidate Second-Order Møller–Plesset Theory with Periodic Boundary Conditions. *J. Chem. Theory Comput.* **2018**, *14* (5), 2427–2438.
CIM	Molecular and nonconducting crystals	Wang, Y.; Ni, Z.; Li, W.; Li, S. Cluster-in-Molecule Local Correlation Approach for Periodic Systems. *J. Chem. Theory Comput.* **2019**, *15* (5), 2933–2943.
iOI	Other (DNA)	Wang, Z.; Liu, W. IOI: An Iterative Orbital Interaction Approach for Solving the Self-Consistent Field Problem. *J. Chem. Theory Comput.* **2021**, *17* (8), 4831–4845.
MOB	Other (peptides)	Saha, A.; Raghavachari, K. Analysis of Different Fragmentation Strategies on a Variety of Large Peptides: Implementation of a Low Level of Theory in Fragment-Based Methods Can Be a Crucial Factor. *J. Chem. Theory*

		Comput. **2015**, *11* (5), 2012–2023.
EE-MTA	Other (proteins)	Isegawa, M.; Wang, B.; Truhlar, D. G. Electrostatically Embedded Molecular Tailoring Approach and Validation for Peptides. *J. Chem. Theory Comput.* **2013**, *9* (3), 1381–1393.
MTA	Other (molecular clusters)	Sahu, N.; Gadre, S. R. Molecular Tailoring Approach: A Route for Ab Initio Treatment of Large Clusters. *Acc. Chem. Res.* **2014**, *47* (9), 2739–2747.
VMB/XPol	Other (molecular clusters)	Gao, J.; Wang, Y. Communication: Variational Many-Body Expansion: Accounting for Exchange Repulsion, Charge Delocalization, and Dispersion in the Fragment-Based Explicit Polarization Method. *J. Chem. Phys.* **2012**, *136* (7), 071101.
EE-MB	Other (molecular clusters)	Qi, H. W.; Leverentz, H. R.; Truhlar, D. G. Water 16-Mers and Hexamers: Assessment of the Three-Body and Electrostatically Embedded Many-Body Approximations of the Correlation Energy or the Nonlocal Energy as Ways to Include Cooperative Effects. *J. Phys. Chem. A* **2013**, *117* (21), 4486–4499.
DFTB	AIMD	Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. Self-Consistent-Charge Density-Functional Tight-Binding Method for Simulations of Complex Materials Properties. *Phys. Rev. B* **1998**, *58* (11), 7260–7268.
DFTB	Fe cluster	Köhler, C.; Seifert, G.; Frauenheim, T. Density Functional Based Calculations for Fe(n), (N≤32). *Chem. Phys.* **2005**, *309* (1), 23–31.
GFN1-xTB	Vibrational spectra	Grimme, S.; Bannwarth, C.; Shushkov, P. A Robust and Accurate Tight-Binding Quantum Chemical Method for Structures, Vibrational Frequencies, and Noncovalent Interactions of Large Molecular Systems Parametrized for All Spd-Block Elements (Z = 1–86). *J. Chem. Theory Comput.* **2017**, *13* (5), 1989–2009.
Excitation energies and excited states		
TD-DFTB	Absorption spectra	Wang, F.; Yam, C. Y.; Chen, G.; Wang, X.; Fan, K.; Niehaus, T. A.; Frauenheim, T. Linear Scaling Time-Dependent Density-Functional Tight-Binding Method for Absorption Spectra of Large Systems. *Phys. Rev. B* **2007**, *76* (4), 45114.
TD-DFTB	Nonadiabatic molecular dynamics	Niehaus, T. A.; Heringer, D.; Torralva, B.; Frauenheim, T. Importance of Electronic Self-Consistency in the TDDFT Based Treatmentof Nonadiabatic Molecular Dynamics. *Eur. Phys. J. D - At. Mol. Opt. Plasma Phys.* **2005**, *35* (3), 467–477.
sTDA-xTB	Electronic spectra	Grimme, S.; Bannwarth, C. Ultra-Fast Computation of Electronic Spectra for Large Systems by Tight-Binding Based Simplified Tamm-Dancoff Approximation (STDA-XTB). *J. Chem. Phys.* **2016**, *145* (5), 54103.
REM	Absorption spectra	Ma, H.; Troisi, A. Direct Optical Generation of Long-Range Charge-Transfer States in Organic Photovoltaics. *Adv. Mater.* **2014**, *26* (35), 6163–6167.
Divide-and-conquer	Absorption spectra	Nakai, H.; Yoshikawa, T. Development of an Excited-State Calculation Method for Large Systems Using Dynamical Polarizability: A Divide-and-Conquer Approach at the Time-Dependent Density Functional Level. *J. Chem. Phys.* **2017**, *146* (12), 124123.
FMO	Absorption spectra	Chiba, M.; Fedorov, D. G.; Kitaura, K. Time-Dependent Density Functional Theory Based upon the Fragment Molecular Orbital Method. *J. Chem. Phys.* **2007**, *127* (10), 104108.
EE-GMFCC	Absorption spectra	Jin, X.; Glover, W. J.; He, X. Fragment Quantum Mechanical Method for Excited States of Proteins: Development and Application to the Green Fluorescent Protein. *J. Chem. Theory Comput.* **2020**, *16* (8), 5174–5188.
Exciton model	Emission spectra	Wang, R.; Yao, Y.; Zhang, C.; Zhang, Y.; Bin, H.; Xue, L.; Zhang, Z.-G.; Xie, X.; Ma, H.; Wang, X.; Li, Y.; Xiao, M. Ultrafast Hole Transfer Mediated by Polaron Pairs in All-Polymer Photovoltaic Blends. *Nat. Commun.* **2019**, *10* (1), 398.
EE-GMF	Aggregation-induced emission	Zhang, W.; Liu, J.; Jin, X.; Gu, X.; Zeng, X. C.; He, X.; Li, H. Quantitative Prediction of Aggregation-Induced Emission: A Full Quantum Mechanical Approach to the Optical Spectra. *Angew. Chemie Int. Ed.* **2020**, *59* (28), 11550–11555.
tDMRG	Absorption and fluorescence spectra	Ren, J.; Shuai, Z.; Kin-Lic Chan, G. Time-Dependent Density Matrix Renormalization Group Algorithms for Nearly Exact Absorption and Fluorescence Spectra of Molecular Aggregates at Both Zero and Finite Temperature. *J. Chem. Theory Comput.* **2018**, *14* (10), 5027–5039.
DMRG	X-ray Raman	Cho, D.; Rouxel, J. R.; Mukamel, S.; Kin-Lic Chan, G.; Li, Z. Stimulated X-Ray Raman and Absorption Spectroscopy of Iron–Sulfur Dimers. *J.*

	spectroscopy (SXRS) and absorption spectroscopy	*Phys. Chem. Lett.* **2019**, *10* (21), 6664–6671.
tDMRG	Two-dimensional ultrafast spectra	Yao, Y.; Sun, K.-W.; Luo, Z.; Ma, H. Full Quantum Dynamics Simulation of a Realistic Molecular System Using the Adaptive Time-Dependent Density Matrix Renormalization Group Method. *J. Phys. Chem. Lett.* **2018**, *9* (2), 413–419.
tDMRG	Time- and Frequency-Resolved Fluorescence	Gelin, M. F.; Borrelli, R. Simulation of Nonlinear Femtosecond Signals at Finite Temperature via a Thermo Field Dynamics-Tensor Train Method: General Theory and Application to Time- and Frequency-Resolved Fluorescence of the Fenna–Matthews–Olson Complex. *J. Chem. Theory Comput.* **2021**, *17* (7), 4316–4331.
FMO	Excitonic Interactions	Kaliakin, D. S.; Nakata, H.; Kim, Y.; Chen, Q.; Fedorov, D. G.; Slipchenko, L. V. FMOxFMO: Elucidating Excitonic Interactions in the Fenna–Matthews–Olson Complex with the Fragment Molecular Orbital Method. *J. Chem. Theory Comput.* **2020**, *16* (2), 1175–1187.
tDMRG	Exciton dissociation	Yao, Y.; Xie, X.; Ma, H. Ultrafast Long-Range Charge Separation in Organic Photovoltaics: Promotion by Off-Diagonal Vibronic Couplings and Entropy Increase. *J. Phys. Chem. Lett.* **2016**, *7* (23), 4830–4835.
tDMRG	Charge transport	Li, W.; Ren, J.; Shuai, Z. A General Charge Transport Picture for Organic Semiconductors with Nonlocal Electron-Phonon Couplings. *Nat. Commun.* **2021**, *12* (1), 4260.
tDMRG	Singlet fission	Mardazad, S.; Xu, Y.; Yang, X.; Grundner, M.; Schollwöck, U.; Ma, H.; Paeckel, S. Quantum Dynamics Simulation of Intramolecular Singlet Fission in Covalently Linked Tetracene Dimer. **2021**, arXiv:2107.13948.
TNS	Singlet fission	Schroder, F. A. Y. N.; Turban, D. H. P.; Musser, A. J.; Hine, N. D. M.; Chin, A. W. Tensor Network Simulation of Multi-Environmental Open Quantum Dynamics via Machine Learning and Entanglement Renormalisation. *Nat. Commun.* **2019**, *10* (1), 1062.
MCTDH	Singlet fission	Tamura, H.; Huix-Rotllant, M.; Burghardt, I.; Olivier, Y.; Beljonne, D. First-Principles Quantum Dynamics of Singlet Fission: Coherent versus Thermally Activated Mechanisms Governed by Molecular π-Stacking. *Phys. Rev. Lett.* **2015**, *115* (10), 107401.
Exciton model	Singlet fission	Morrison, A. F.; Herbert, J. M. Evidence for Singlet Fission Driven by Vibronic Coherence in Crystalline Tetracene. *J. Phys. Chem. Lett.* **2017**, *8* (7), 1442–1448.
Exciton model	Singlet fission	Wang, Z.; Liu, H.; Xie, X.; Zhang, C.; Wang, R.; Chen, L.; Xu, Y.; Ma, H.; Fang, W.; Yao, Y.; Sang, H.; Wang, X.; Li, X.; Xiao, M. Free-Triplet Generation with Improved Efficiency in Tetracene Oligomers through Spatially Separated Triplet Pair States. *Nat. Chem.* **2021**, *13* (6), 559–567.
DMRG	Singlet fission	Ren, J.; Peng, Q.; Zhang, X.; Yi, Y.; Shuai, Z. Role of the Dark 2Ag State in Donor–Acceptor Copolymers as a Pathway for Singlet Fission: A DMRG Study. *J. Phys. Chem. Lett.* **2017**, *8* (10), 2175–2181.
MCTDH	Real-time nonadiabatic dynamics	Tamura, H.; Huix-Rotllant, M.; Burghardt, I.; Olivier, Y.; Beljonne, D. First-Principles Quantum Dynamics of Singlet Fission: Coherent versus Thermally Activated Mechanisms Governed by Molecular π-Stacking. *Phys. Rev. Lett.* **2015**, *115* (10), 107401.
DMRG	$[Fe_2S_2]$ and $[Fe_4S_4]$ clusters	Sharma, S.; Sivalingam, K.; Neese, F.; Chan, G. K.-L. Low-Energy Spectrum of Iron–Sulfur Clusters Directly from Many-Particle Quantum Mechanics. *Nat. Chem.* **2014**, *6* (10), 927–933.
DMRG	$[Fe_8S_7]$ P-cluster	Li, Z.; Guo, S.; Sun, Q.; Chan, G. K.-L. Electronic Landscape of the P-Cluster of Nitrogenase as Revealed through Many-Electron Quantum Wavefunction Simulations. *Nat. Chem.* **2019**, *11* (11), 1026–1033.
DMRG	Mn_4CaO_5	Kurashige, Y.; Chan, G. K.-L.; Yanai, T. Entangled Quantum Electronic Wavefunctions of the Mn4CaO5 Cluster in Photosystem II. *Nat. Chem.* **2013**, *5* (8), 660–666.

Table S2. Selected works of using machine learning to study materials.

ML method	Systems/Properties	Reference
Random forest	Organic photovoltaic	Sun, W.; Zheng, Y.; Yang, K.; Zhang, Q.; Shah, A. A.; Wu, Z.; Sun, Y.; Feng, L.; Chen, D.; Xiao, Z.; Lu, S.; Li, Y.; Sun, K. Machine Learning-Assisted Molecular Design and Efficiency Prediction for High-Performance Organic Photovoltaic Materials. *Sci. Adv.* **2019**, *5* (11). eaay4275.
Neural network	Organic photovoltaic	Sun, W.; Li, M.; Li, Y.; Wu, Z.; Sun, Y.; Lu, S.; Xiao, Z.; Zhao, B.; Sun, K. The Use of Deep Learning to Fast Evaluate Organic Photovoltaic Materials. *Adv. Theory Simulations* **2019**, *2* (1), 1800116.
Kernel ridge regression	Organic photovoltaic	Padula, D.; Troisi, A. Concurrent Optimization of Organic Donor–Acceptor Pairs through Machine Learning. *Adv. Energy Mater.* **2019**, *9* (40), 1902463.
Gradient boosting regression tree	Organic photovoltaic	Sahu, H.; Yang, F.; Ye, X.; Ma, J.; Fang, W.; Ma, H. Designing Promising Molecules for Organic Solar Cells via Machine Learning Assisted Virtual Screening. *J. Mater. Chem. A* **2019**, *7* (29), 17480–17488.
Gradient boosting regression tree	Organic photovoltaic	Sahu, H.; Ma, H. Unraveling Correlations between Molecular Properties and Device Parameters of Organic Solar Cells Using Machine Learning. *J. Phys. Chem. Lett.* **2019**, *10* (22), 7277–7284.
Kernel ridge regression	Organic photovoltaic	Padula, D.; Simpson, J. D.; Troisi, A. Combining Electronic and Structural Features in Machine Learning Models to Predict Organic Solar Cells Properties. *Mater. Horiz.* **2019**, *6* (2), 343–349.
Gradient boosting regression tree	Organic photovoltaic	Sahu, H.; Rao, W.; Troisi, A.; Ma, H. Toward Predicting Efficiency of Organic Solar Cells via Machine Learning and Improved Descriptors. *Adv. Energy Mater.* **2018**, *8* (24), 1801032.
Random forest	Organic photovoltaic	Nagasawa, S.; Al-Naamani, E.; Saeki, A. Computer-Aided Screening of Conjugated Polymers for Organic Solar Cell: Classification by Random Forest. *J. Phys. Chem. Lett.* **2018**, *9* (10), 2639–2646.
Gradient Boosting Regression Tree (GB)	organic photovoltaics (OPVs)	Sahu, H.; Rao, W.; Troisi, A.; Ma, H. Toward Predicting Efficiency of Organic Solar Cells via Machine Learning and Improved Descriptors. *Adv. Energy Mater.* **2018**, *8* (24), 1801032.
Neural network	Organic molecules	Pyzer-Knapp, E. O.; Li, K.; Aspuru-Guzik, A. Learning from the Harvard Clean Energy Project: The Use of Neural Networks to Accelerate Materials Discovery. *Adv. Funct. Mater.* **2015**, *25* (41), 6495–6502.
Random forest	Organic light-emitting diodes	Lee, M.-H. Identification of Host–Guest Systems in Green TADF-Based OLEDs with Energy Level Matching Based on a Machine-Learning Study. *Phys. Chem. Chem. Phys.* **2020**, *22* (28), 16378–16386.
Random forest	Organic field-effect transistors	Lee, M.-H. Machine Learning for Understanding the Relationship between the Charge Transport Mobility and Electronic Energy Levels for N-Type Organic Field-Effect Transistors. *Adv. Electron. Mater.* **2019**, *5* (12), 1900573.
Random forest	Organic light-emitting diodes	Janai, M. A. Bin; Woon, K. L.; Chan, C. S. Design of Efficient Blue Phosphorescent Bottom Emitting Light Emitting Diodes by Machine Learning Approach. *Org. Electron.* **2018**, *63*, 257–266.
Neural network	Organic light-emitting diodes (OLEDs)	R. Gómez-Bombarelli, J. Aguilera-Iparraguirre, T. D. Hirzel, D. Duvenaud, D. Maclaurin, M. A. Blood-Forsythe, H. S. Chae, M. Einzinger, D.-G. Ha, T. Wu, G. Markopoulos, S. Jeon, H. Kang, H. Miyazaki, M. Numata, S. Kim, W. Huang, S. I. Hong, M. Baldo, R. P. Adams, A. Aspuru-Guzik, *Nat. Mater.*, 2016, **15**, 1120-1128.
Neural network (Generative model)	Organic redox flow batteries	Sanchez-Lengeling, B.; Aspuru-Guzik, A. Inverse Molecular Design Using Machine Learning: Generative Models for Matter Engineering. *Science,* **2018**, *361* (6400), 360–365.
Kernel ridge regression, support vector machine and neural network	Porphyrins	Li, Z.; Omidvar, N.; Chin, W. S.; Robb, E.; Morris, A.; Achenie, L.; Xin, H. Machine-Learning Energy Gaps of Porphyrins with Molecular Graph Representations. *J. Phys. Chem. A* **2018**, *122* (18), 4571–4578.
Density-based clustering	Supramolecular Polymers	Gasparotto, P.; Bochicchio, D.; Ceriotti, M.; Pavan, G. M. Identifying and Tracking Defects in Dynamic Supramolecular Polymers. *J. Phys. Chem. B* **2020**, *124* (3), 589–599.
Uniform Manifold Approximation and Projection	Conductive polymers	Magdau, I.-B.; Miller, T. F. Machine Learning Solvation Environments in Conductive Polymers: Application to ProDOT-2Hex with Solvent Swelling. *Macromolecules* **2021**, *54* (7), 3377–3387.

Neural Network	High thermal conductivity polymers	Zhu, M.-X.; Song, H.-G.; Yu, Q.-C.; Chen, J.-M.; Zhang, H.-Y. Machine-Learning-Driven Discovery of Polymers Molecular Structures with High Thermal Conductivity. *Int. J. Heat Mass Transf.* **2020**, *162*, 120381.
Transfer Learning	High thermal conductivity polymers	Wu, S.; Kondo, Y.; Kakimoto, M.; Yang, B.; Yamada, H.; Kuwajima, I.; Lambard, G.; Hongo, K.; Xu, Y.; Shiomi, J.; Schick, C.; Morikawa, J.; Yoshida, R. Machine-Learning-Assisted Discovery of Polymers with High Thermal Conductivity Using a Molecular Design Algorithm. *npj Comput. Mater.* **2019**, *5* (1), 66.
Neural network	Thermoelectric materials	Iwasaki, Y.; Takeuchi, I.; Stanev, V.; Kusne, A. G.; Ishida, M.; Kirihara, A.; Ihara, K.; Sawada, R.; Terashima, K.; Someya, H.; Uchida, K.; Saitoh, E.; Yorozu, S. Machine-Learning Guided Discovery of a New Thermoelectric Material. *Sci. Rep.* **2019**, *9* (1), 2751.
Gaussian Process	Crystalline polymers	Hong, S. J.; Chun, H.; Lee, J.; Kim, B.-H.; Seo, M. H.; Kang, J.; Han, B. First-Principles-Based Machine-Learning Molecular Dynamics for Crystalline Polymers with van Der Waals Interactions. *J. Phys. Chem. Lett.* **2021**, *12* (25), 6000–6006.
Neural Network	Polymer	St. John, P. C.; Phillips, C.; Kemper, T. W.; Wilson, A. N.; Guan, Y.; Crowley, M. F.; Nimlos, M. R.; Larsen, R. E. Message-Passing Neural Networks for High-Throughput Polymer Screening. *J. Chem. Phys.* **2019**, *150* (23), 234111.
Neural Network	Polymer	Doan Tran, H.; Kim, C.; Chen, L.; Chandrasekaran, A.; Batra, R.; Venkatram, S.; Kamal, D.; Lightstone, J. P.; Gurnani, R.; Shetty, P.; Ramprasad, M.; Laws, J.; Shelton, M.; Ramprasad, R. Machine-Learning Predictions of Polymer Properties with Polymer Genome. *J. Appl. Phys.* **2020**, *128* (17), 171104.
Gaussian Process	Polyacrylamides	Zhang, Y.; Xu, X. Machine Learning Glass Transition Temperature of Polyacrylamides Using Quantum Chemical Descriptors. *Polym. Chem.* **2021**, *12* (6), 843–851.
Gaussian process	Organic semiconductors	Yang, J.; De, S.; Campbell, J. E.; Li, S.; Ceriotti, M.; Day, G. M. Large-Scale Computational Screening of Molecular Organic Semiconductors Using Crystal Structure Prediction. *Chem. Mater.* **2018**, *30* (13), 4361–4371.
Neural Network	Organic semiconductors	Lu, C.; Liu, Q.; Sun, Q.; Hsieh, C.-Y.; Zhang, S.; Shi, L.; Lee, C.-K. Deep Learning for Optoelectronic Properties of Organic Semiconductors. *J. Phys. Chem. C* **2020**, *124* (13), 7048–7060.
Solid materials		
Gaussian Process	Molecular crystals	Musil, F.; De, S.; Yang, J.; Campbell, J. E.; Day, G. M.; Ceriotti, M. Machine Learning for the Structure–Energy–Property Landscapes of Molecular Crystals. *Chem. Sci.* **2018**, *9* (5), 1289–1300.
Gaussian process	Molecular crystals	McDonagh, D.; Skylaris, C. K.; Day, G. M. Machine-Learned Fragment-Based Energies for Crystal Structure Prediction. *J. Chem. Theory Comput.* **2019**, *15* (4), 2743–2758.
Gaussian Process	Molecular crystals	Raimbault, N.; Grisafi, A.; Ceriotti, M.; Rossi, M. Using Gaussian Process Regression to Simulate the Vibrational Raman Spectra of Molecular Crystals. *New J. Phys.* **2019**, *21* (10), 105001.
Gaussian Process	Molecular Crystal	Wengert, S.; Csányi, G.; Reuter, K.; Margraf, J. T. Data-Efficient Machine Learning for Molecular Crystal Structure Prediction. *Chem. Sci.* **2021**, *12* (12), 4536–4546.
Extreme gradient boosting and random forest	Zeolite	Muraoka, K.; Sada, Y.; Miyazaki, D.; Chaikittisilp, W.; Okubo, T. Linking Synthesis and Structure Descriptors from a Large Collection of Synthetic Records of Zeolite Materials. *Nat. Commun.* **2019**, *10* (1), 4459.
Gradient boosting regression tree	Zeolite	Evans, J. D.; Coudert, F.-X. Predicting the Mechanical Properties of Zeolite Frameworks by Machine Learning. *Chem. Mater.* **2017**, *29* (18), 7833–7839.
Extreme gradient boosting	Zeolite	Gu, Y.; Liu, Z.; Yu, C.; Gu, X.; Xu, L.; Gao, Y.; Ma, J. Zeolite Adsorption Isotherms Predicted by Pore Channel and Local Environmental Descriptors: Feature Learning on DFT Binding Strength. *J. Phys. Chem. C* **2020**, *124* (17), 9314–9328.
Neural Network	Zeolite	Daeyaert, F.; Ye, F.; Deem, M. W. Machine-Learning Approach to the Design of OSDAs for Zeolite Beta. *Proc. Natl. Acad. Sci. USA* **2019**, *116* (9), 3413–3418.
Random forest	Zeolite	Carr, D. A.; Lach-hab, M.; Yang, S.; Vaisman, I. I.; Blaisten-Barojas, E. Machine Learning Approach for Structure-Based Zeolite Classification. *Microporous Mesoporous Mater.* **2009**, *117* (1), 339–349.
Decision trees	Zeolite	Huo, W.; Gao, N.; Yan, Y.; Li, J.; Yu, J.; Xu, R. Decision Trees Combined

		with Feature Selection for the Rational Synthesis of Aluminophosphate $AlPO_4$-5. *Acta Physico-Chimica Sin.* **2011**, *27* (09), 2111.
Random forest and support vector machine	Metal-organic framework materials	Batra, R.; Chen, C.; Evans, T. G.; Walton, K. S.; Ramprasad, R. Prediction of Water Stability of Metal–Organic Frameworks Using Machine Learning. *Nat. Mach. Intell.* **2020**, *2* (11), 704–710.
Neural network (Meta learning)	MOFs	Sun, Y.; DeJaco, R. F.; Li, Z.; Tang, D.; Glante, S.; Sholl, D. S.; Colina, C. M.; Snurr, R. Q.; Thommes, M.; Hartmann, M.; Siepmann, J. I. Fingerprinting Diverse Nanoporous Materials for Optimal Hydrogen Storage Conditions Using Meta-Learning. *Sci. Adv.* **2021**, *7* (30), eabg3983.
Neural Network	Metal-organic framework materials	Eckhoff, M.; Behler, J. From Molecular Fragments to the Bulk: Development of a Neural Network Potential for MOF-5. *J. Chem. Theory Comput.* **2019**, *15* (6), 3793–3809.
Random forest	Metal-organic framework materials	Anderson, R.; Rodgers, J.; Argueta, E.; Biong, A.; Gómez-Gualdrón, D. A. Role of Pore Chemistry and Topology in the CO2 Capture Capabilities of MOFs: From Molecular Simulation to Machine Learning. *Chem. Mater.* **2018**, *30* (18), 6325–6337.
Neural Network	Metal-organic framework materials	Raza, A.; Sturluson, A.; Simon, C. M.; Fern, X. Message Passing Neural Networks for Partial Charge Assignment to Metal–Organic Frameworks. *J. Phys. Chem. C* **2020**, *124* (35), 19070–19082.
Transfer Learning	Metal-organic framework materials	Ma, R.; Colón, Y. J.; Luo, T. Transfer Learning Study of Gas Adsorption in Metal–Organic Frameworks. *ACS Appl. Mater. & Interfaces* **2020**, *12* (30), 34041–34048.
Gradient Boosting Decision Tree (GBDT)	Inorganic crystals	Isayev, O.; Oses, C.; Toher, C.; Gossett, E.; Curtarolo, S.; Tropsha, A. Universal Fragment Descriptors for Predicting Properties of Inorganic Crystals. *Nat. Commun.* **2017**, *8* (1), 15679.
LASSO (least absolute shrinkage and selection operator	Metal nanoparticles	Dean, J.; Taylor, M. G.; Mpourmpakis, G. Unfolding Adsorption on Metal Nanoparticles: Connecting Stability with Catalysis. *Sci. Adv.* **2019**, *5* (9), eaax5101.
Active Machine Learning	Electrocatalysts	Zhong, M.; Tran, K.; Min, Y.; Wang, C.; Wang, Z.; Dinh, C.-T.; De Luna, P.; Yu, Z.; Rasouli, A. S.; Brodersen, P.; Sun, S.; Voznyy, O.; Tan, C.-S.; Askerka, M.; Che, F.; Liu, M.; Seifitokaldani, A.; Pang, Y.; Lo, S.-C.; Ip, A.; Ulissi, Z.; Sargent, E. H. Accelerated Discovery of CO2 Electrocatalysts Using Active Machine Learning. *Nature* **2020**, *581* (7807), 178–183.
Gradient Boosting Decision Tree (GBDT)	hybrid organic-inorganic perovskites (HOIPs)	Lu, S.; Zhou, Q.; Ouyang, Y.; Guo, Y.; Li, Q.; Wang, J. Accelerated Discovery of Stable Lead-Free Hybrid Organic-Inorganic Perovskites via Machine Learning. *Nat. Commun.* **2018**, *9* (1), 3405.
Other systems		
Graph convolutional Neural network	Hydroxyapatite Nanoparticles (HANPs)	Liu, Z.; Shi, Y.; Chen, H.; Qin, T.; Zhou, X.; Huo, J.; Dong, H.; Yang, X.; Zhu, X.; Chen, X.; Zhang, L.; Yang, M.; Gao, Y.; Ma, J. Machine Learning on Properties of Multiscale Multisource Hydroxyapatite Nanoparticles Datasets with Different Morphologies and Sizes. *npj Comput. Mater.* **2021**, *7* (1), 142.
Gaussian Process	Fluid methane	Veit, M.; Jain, S. K.; Bonakala, S.; Rudra, I.; Hohl, D.; Csányi, G. Equation of State of Fluid Methane from First Principles with Machine Learning Potentials. *J. Chem. Theory Comput.* **2019**, *15* (4), 2574–2586.
Neural Network	Liquid and solid water	Cheng, B.; Engel, E. A.; Behler, J.; Dellago, C.; Ceriotti, M. Ab Initio Thermodynamics of Liquid and Solid Water. *Proc. Natl. Acad. Sci. USA* **2019**, *116* (4), 1110–1115.
Neural Network	Ionic liquid	Low, K.; Kobayashi, R.; Izgorodina, E. I. The Effect of Descriptor Choice in Machine Learning Models for Ionic Liquid Melting Point Prediction. *J. Chem. Phys.* **2020**, *153* (10), 104101
t-distributed stochastic neighbor embedding method	Glassy liquid	Cubuk, E. D.; Schoenholz, S. S.; Kaxiras, E.; Liu, A. J. Structural Properties of Defects in Glassy Liquids. *J. Phys. Chem. B* **2016**, *120* (26), 6139–6146.