This is an open access article published under a Creative Commons Non-Commercial No Derivative Works (CC-BY-NC-ND) Attribution License, which permits copying and redistribution of the article, and creation of adaptations, all for non-commercial purposes.

JCTC Journal of Chemical Theory and Computation

Article

Cite This: J. Chem. Theory Comput. 2019, 15, 1777−1784

pubs.acs.org/JCTC

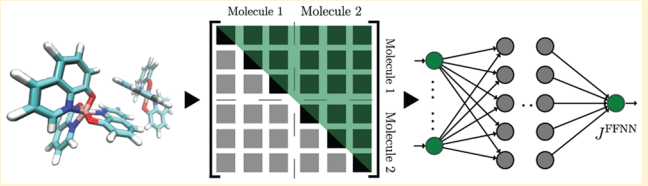
Evolutionary Approach to Constructing a Deep Feedforward Neural Network for Prediction of Electronic Coupling Elements in Molecular Materials

Onur Çaylak,*,† Anil Yaman,*,‡ and Björn Baumeier*,‡

†Department of Mathematics and Computer Science & Institute for Complex Molecular Systems, Eindhoven University of Technology, P.O. Box 513, 5600MB Eindhoven, The Netherlands

‡Department of Mathematics and Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600MB Eindhoven, The Netherlands

ABSTRACT: We present a general framework for the construction of a deep feedforward neural network (FFNN) to predict distance and orientation dependent electronic coupling elements in disordered molecular materials. An evolutionary algorithm automatizes the selection of an optimal architecture of the artificial neural network within a predefined search space. Systematic guidance, beyond minimizing the model error with stochastic gradient descent based backpropagation, is provided by simultaneous maximization of a model fitness that takes into account additional physical properties, such as the field-dependent carrier mobility. As a prototypical system, we consider hole transport in amorphous tris(8-hydroxyquinolinato)aluminum. Reference data for training and validation is obtained from multiscale ab initio simulations, in which coupling elements are evaluated using density-functional theory, for a system containing 4096 molecules. The Coulomb matrix representation is chosen to encode the explicit molecular pair coordinates into a rotation and translation invariant feature set for the FFNN. The final optimized deep feedforward neural network is tested for transport models without and with energetic disorder. It predicts electronic coupling elements and mobilities in excellent agreement with the reference data. Such a FFNN is readily applicable to much larger systems at negligible computational cost, providing a powerful surrogate model to overcome the size limitations of the ab initio approach.



1. INTRODUCTION

Dynamics of electronic excitations drives the functionality of molecular nanomaterials in many energy applications, e.g., in organic photovoltaics, photocatalysis, thermoelectricity, or energy storage. The dynamics is governed not only by the chemical structure, architecture, or electronic structure of molecular building blocks but also by the local and global morphology of the materials and molecular interactions on the mesoscale.[1−3] It is essential to understand how elementary dynamic processes, such as electron or energy transfer, emerge from an interplay of morphology and electronic structure. Such fundamental insight will eventually allow for controlling the above processes and pave the way for a rational design of molecular materials. While macroscale information can be obtained experimentally, zooming into the electronic dynamics at an (sub)atomic scale is nearly impossible.[4]

Computational modeling of, e.g., charge dynamics can provide valuable insight in this situation. The Gaussian disorder model and its various extensions[5−8] have been used to study general aspects of transport, such as temperature or carrier density dependence.[9,10] For material specificity, they require information about the width of the density of states, which is typically obtained by fitting to macroscale or device-scale observables, for instance, a current−voltage curve. These more descriptive models cannot provide detailed information about underlying intermolecular processes.

In contrast, bottom-up simulations of charge and exciton dynamics in large-scale morphologies aim to explicitly zoom in to elementary charge transfer reactions at molecular level and predict the mesoscale charge dynamics using multiscale strategies which link quantum and supramolecular scales.[11−13] Such approaches have shown a remarkable level of predictiveness[3,14−16] but come at the price of high computational costs. They typically involve the determination of bimolecular electron transfer rates in explicit material morphologies, which in turn requires the calculation of intermolecular electronic coupling elements,[17,18] or transfer integrals, of the form

$$J_{ij} = \langle \Psi_i | \hat{H} | \Psi_j \rangle \tag{1}$$

where $|\Psi_i\rangle$ and $|\Psi_j\rangle$ are diabatic states of molecules i and j, respectively, and $\hat{H}$ is the Hamiltonian of the coupled system. Practical evaluation of eq 1 relies on quantum-mechanical information about the relevant electronic states of the two individual monomers, as well as of the dimer formed by them.

Received: December 21, 2018
Published: February 12, 2019

© 2019 American Chemical Society

DOI: 10.1021/acs.jctc.8b01285
J. Chem. Theory Comput. 2019, 15, 1777−1784

Based on density-functional theory (DFT) the necessary calculations for a morphology consisting of a few thousand molecules of moderate size can consume several hundreds of days of CPU time, even with techniques optimized for large-scale systems. Many relevant materials or processes, e.g., the system-size dependence of carrier mobility in dispersive transport, realistic carrier densities, or disordered interfaces in heterojunctions, can only be studied using significantly larger systems that are currently not accessible to multiscale models.

Surrogate stochastic models have been developed to overcome some of these limitations.[19,20] They represent the molecular morphology by (random) point patterns and assign coupling elements between them by drawing from distributions with distance-dependent means and standard deviations, fitted to microscopic data. These models manage to reproduce, e.g., the field-dependence of the mobility stochastically, i.e., averages over several realizations are required to obtain a comparable behavior. While the generation of larger surrogate systems is computationally inexpensive, information about the molecular details is lost, and the models are not transferable to interfaces or heterostructures.

In this paper, we develop an alternative surrogate model which allows application to system sizes currently inaccessible to the multiscale ab initio approach while retaining its molecular level details. Focus is on removing the computational bottleneck associated with the explicit quantum-mechanical evaluations of electronic couplings using eq 1 with the help of Machine Learning (ML). ML has attracted considerable interest as a tool to save computational costs in large scale calculations[21] or in exploring chemical space, e.g., by predicting material properties.[22−25] Different models differ by the representation of the molecular information (features) and the choice of a suitable ML algorithm, and their combinations have been optimized accordingly.[23,26]

For our goal of predicting electronic coupling elements, an appropriate data representation must accurately take distance and mutual orientations between two molecules of the same type, as given by the explicit atomic coordinates, into account. The machine learning algorithm must be capable of reliably predicting values of J_{ij} that can easily span several orders of magnitude, in particular in amorphous molecular materials. In this situation we turn to biologically inspired computational models known as artificial neural networks (ANNs).[27] However, the construction of an appropriate network architecture is not trivial and may require a trial-and-error approach. We deal with this problem in a systematic way by using search algorithms such as evolutionary algorithms (EA).[28] The advantage of using an EA approach to constructing a neural network is that it not only minimizes the model error but is also capable of taking into account additional physical principles providing systematic guidance to designing architectures.

Specifically, we employ such an evolutionary method to construct a multilayered (deep) feedforward neural network (FFNN) for the prediction of electronic coupling elements based on the Coulomb Matrix representation[26] of molecular orientations. As a prototypical system, we consider an amorphous morphology of tris(8-hydroxyquinolinato)aluminum (Alq_3). An ab initio model of hole transport explicitly determined for a system containing 4096 molecules serves as a reference for the training of the FFNN. The electric-field dependence of the hole mobility is used as an additional fitness parameter in the evolutionary algorithm. The final FFNN model provides inexpensive predictions of J_{ij} with which hole mobilities are obtained in excellent agreement with the ab initio data, both without and with energetic disorder. We demonstrate that it is readily applicable to systems of larger size containing 8192 and 32768 molecules, respectively.

This paper is organized as follows: In section 2, we briefly recapitulate steps in the multiscale ab initio model to obtain the reference data. Details about the data representation and processing and the evolutionary approach for constructing the feedforward neural network including the definition of its fitness are given in section 3. Validation of the model and prediction results are discussed in section 4. A brief summary concludes the paper.

2. MULTISCALE AB INITIO MODEL

In what follows, we briefly summarize the steps of the multiscale model of charge transport in disordered molecular materials, needed to generate the ab initio reference for the FFNN. More in-depth discussion of the procedures can be found in ref 11. The starting point is the simulation of an atomistic morphology using classical Molecular Dynamics (MD). 4096 Alq_3 molecules are arranged randomly in a cubic box, which is then first equilibrated above the glass transition temperature in an NpT ensemble at T = 700 K and p = 1 bar and subsequently annealed to T = 300 K. The Berendsen barostat[29] with a time constant of 2.0 ps and the velocity rescaling thermostat[30] with a time constant of 0.5 ps are used throughout. All calculations make use of an OPLS-based force field specifically developed[31] for Alq_3 and are performed using Gromacs.[32]

In the obtained room-temperature morphology, the centers of mass of the molecules define the hopping sites for charge carriers. Pairs of molecules for which any intermolecular distance of the respective 8-hydroxyquinoline ligands is less than 0.8 nm are added to a neighborlist of possible charge transfer pairs. Transfer rates between two molecules i and j are calculated within the high-temperature limit of nonadiabatic transfer theory[33] using the Marcus expression, which is given by

$$\omega_{ij} = \frac{2\pi}{\hbar}\frac{J_{ij}^2}{\sqrt{4\pi\lambda_{ij}k_B T}}\exp\left[-\frac{(\Delta G_{ij} - \lambda_{ij})^2}{4\lambda_{ij}k_B T}\right] \quad (2)$$

where $\hbar$ is the reduced Planck constant, T is the temperature, λ_{ij} is the reorganization energy, k_B is the Boltzmann's constant, ΔG_{ij} is the free energy difference between initial and final states, and J_{ij} is the electronic coupling element, as defined in eq 1. A hole reorganization energy of 0.23 eV is obtained from DFT calculations with the B3LYP functional[34] and a triple-ζ basis set.[35] Site energies E_i are determined from a classical atomistic model that takes into account the effects of local electric fields and polarization in the morphology.[36,37] Their distribution is approximately Gaussian with σ = 0.20 eV. The influence of an externally applied electric field $\mathbf{F}$ is added to the site energy difference as $\Delta G_{ij} = \Delta E_{ij} + \Delta E_{ij}^{ext} = \Delta E_{ij} + e\mathbf{F}\mathbf{r}_{ij}$, where $\mathbf{r}_{ij}$ is the distance vector between molecules i and j. Electronic coupling elements are calculated using a dimer-projection technique based on DFT[18] with the Perdew−Burke−Ernzerhof functional[38] and the triple-ζ basis. All DFT calculations are performed with the Orca software package,[39] while the VOTCA package[11,40] is used for all charge transport related steps.

DOI: 10.1021/acs.jctc.8b01285
J. Chem. Theory Comput. 2019, 15, 1777−1784

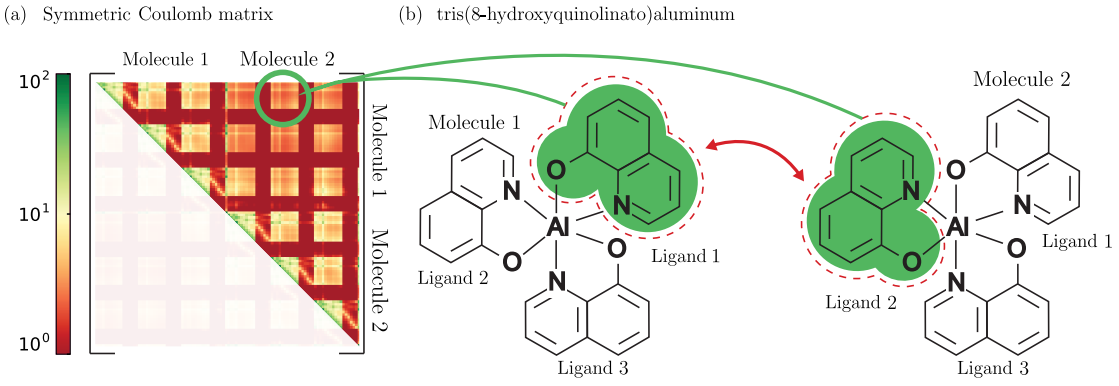


Figure 1. Visualization of a Coulomb matrix for molecular pairs. (a) The Coulomb matrix contains values corresponding to the inter- and intramolecular electrostatic interactions. Lowest values (dark red) correspond to the relations of hydrogen atoms, whereas interactions among heavy atoms of the ligands lead to values near 10 (yellow). (b) Schematic representation of how the arrangement of two Alq_3 molecules is encoded in specific regions of the upper right block of the CM.

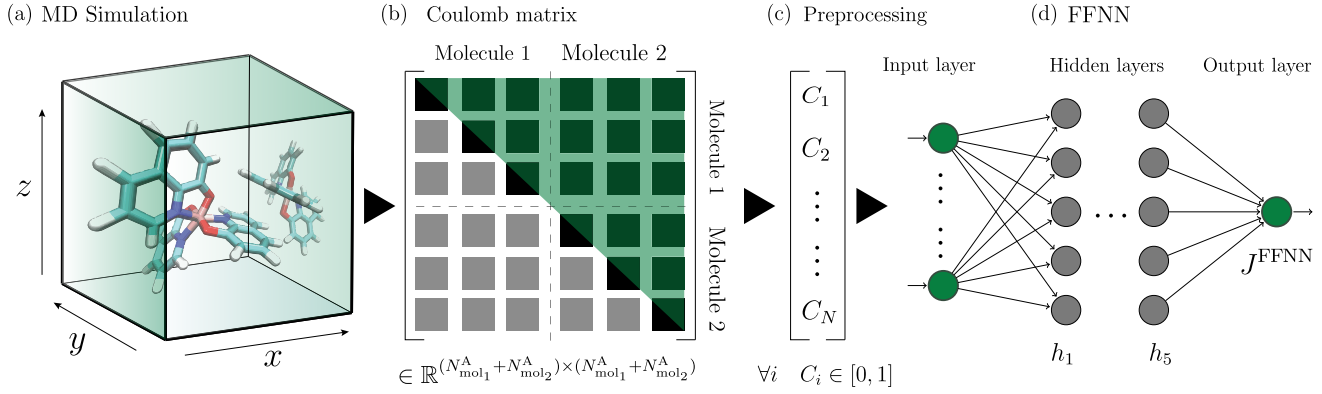


Figure 2. Overview of the data flow from raw molecular dynamics information to a neural network input. (a) Explicit atomic coordinates of molecular pairs is extracted from an MD snapshot. (b) A symmetric Coulomb matrix with dimension given by the sum of the number of atoms per molecule $(N^A_{\mathrm{mol}_1} + N^A_{\mathrm{mol}_2})^2$ is constructed. (c) To only keep relevant and nonredundant information in the vectorized form of the Coulomb matrix, preprocessing techniques such as feature selection (upper triangle) and data scaling (to [0,1]) are introduced. (d) The final vectorized CM enters a feedforward neural network with the hidden layers $h_1,\ldots,h_5$ to predict the electronic coupling elements J^{FFNN}_{ij}.

The molecular centers of mass and the hopping rates between the molecules can be interpreted as the vertices and edges of a weighted directed graph. In a system with only one charge carrier, the time evolution of the occupation probabilities P_i is described by the Kolmogorov forward equation (Master equation)

$$\frac{\mathrm{d}P_i}{\mathrm{d}t} = \sum_j P_j\omega_{ji} - \sum_j P_i\omega_{ij} \tag{3}$$

However, in this work, we are interested in a system that is in a steady state. This restriction allows us to write eq 3 as

$$\sum_j P_j\omega_{ji} - \sum_j P_i\omega_{ij} = 0 \Rightarrow \mathbf{Wp} = 0 \tag{4}$$

Here, the matrix $\mathbf{W}$ can be constructed from the Marcus rates ω_{ij}. The field-dependent mobility $\mu(\mathbf{F})$ can be obtained from steady state occupation probabilities via the relation

$$\langle \mathbf{v} \rangle = \sum_i \sum_j P_j\omega_{ji}(\mathbf{r}_i - \mathbf{r}_j) = \mu(\mathbf{F})\mathbf{F} \tag{5}$$

where $\langle \mathbf{v} \rangle$ is the average velocity.

3. MACHINE LEARNING MODEL

3.1. Data Representation. Explicit structural information on molecular pair geometries is extracted from MD simulations and used to construct the features of the data set. Featurization is the process of encoding molecules into vectors, where each vector gets a label, in this case $\log_{10}[(J_{ij}/\mathrm{eV})^2]$.

Coupling elements between molecular pairs are translation and rotation invariant, which is not accounted for in the plain atom coordinates $\mathbf{R}_i$. The Coulomb matrix (CM)[23,26,41] representation is capable of capturing this invariance and is used in the following to encode crucial information into the data set's features.

For every molecular pair the entries c_{ij} of the corresponding Coulomb matrix $\mathbf{C}$ are computed according to

$$c_{ij} := \begin{cases} \frac{1}{2}Z_i^{2.4} & \text{if } i = j \\ \frac{Z_iZ_j}{\|\mathbf{R}_i - \mathbf{R}_j\|} & \text{if } i \neq j \end{cases} \tag{6}$$

where Z_i is the nuclear charge of atom i. Figure 1 illustrates the building blocks of the CM representation applied to a pair of molecules. One obtains a symmetric matrix that consists of four equally sized block matrices. The upper-right block

DOI: 10.1021/acs.jctc.8b01285
J. Chem. Theory Comput. 2019, 15, 1777−1784

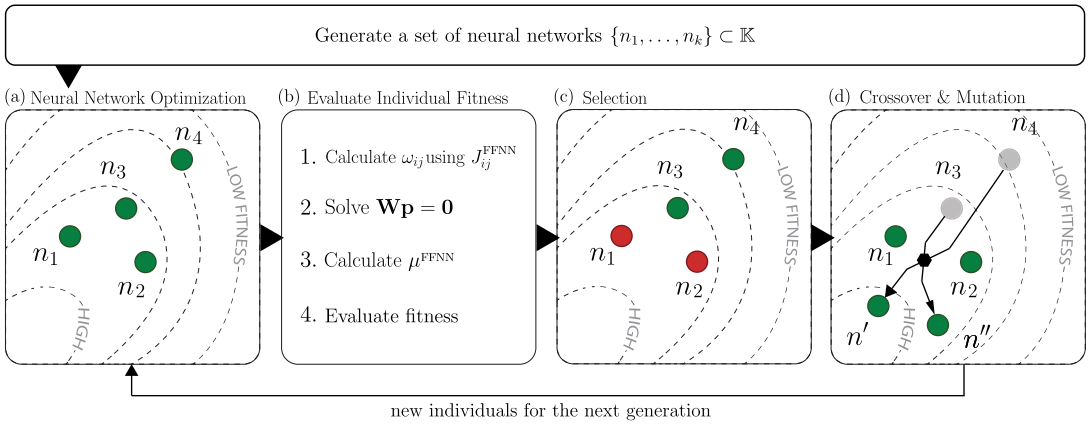


Figure 3. Flowchart of the evolutionary algorithm used for optimizing the deep FFNN. The initialization process consists of generating k = 30 arbitrary FFNN architectures $\{n_1,...,n_k\}$. (a) The weights of the networks are then optimized with a backpropagation algorithm. (b) For every optimized network n_i the predicted electronic coupling elements J_{ij}^{FFNN} are used in eq 2 to determine the matrix **W**. After solving the stationary Master eq (eq 4) and calculation of the field-dependent mobility, the fitness of the architectures is evaluated. (c) The architectures are then ordered based on their fitnesses and selected according to the roulette wheel principle. (d) After applying the crossover and mutation operators, a new generation of feedforward neural networks is generated, and the whole process is repeated until one of the stopping criteria is satisfied.

indicates the intermolecular orientations, whereas the upper-left and lower-right matrices represent the intramolecular conformations.

Before being usable as input for the FFNN, the calculated CMs must be preprocessed, as illustrated in Figure 2(b,c). This step involves the removal of the lower triangular entries including the diagonal elements and scaling of the values to the interval $[0,1]$. Subsequent vectorization yields the instances of the descriptor space as input to an artificial neural network.

3.2. Deep Feedforward Neural Networks and Evolutionary Algorithms. Artificial neural networks consist of a number of artificial neurons typically arranged as layers with specific connectivity referred to as topology. Among the ANNs, feed forward neural networks arrange a certain number of consecutive layers of neurons where each neuron in each layer is directionally connected to all neurons within the next layer. The activation a_i^m of neuron i in layer m is computed from the activations of neurons in the $m-1$-th layer according to

$$a_i^m = f\left(\sum_{j=0} \nu_{ij}^{m-1,m} \cdot a_j^{m-1}\right) \tag{7}$$

where $\nu_{i,j}$ is the weight of the connection between the neurons, and f is an activation functions. For our purposes this activation function is given by a sigmoid function.

One of the conventional ways of training the FFNNs is the backpropagation algorithm with stochastic gradient descent. However, the number of layers and the number of neurons per each layer should be defined before the training. These parameters are referred to as hyperparameters and play an important role in the performance of the networks. Although there are some "rule of thumb" guidelines established based on empirical studies, the selection of a proper set of hyperparameter settings may require a great deal of expert knowledge and/or trial-and-error. This can be avoided by search algorithms like Genetic Algorithms (GAs). GAs are population based global search algorithms inspired by biological evolution.[28] The research field known as Neuroevolution employs evolutionary computing approaches to optimize artificial neural networks.[42] Adopting the terminology from biology, the genetic material of a population of individuals encodes parameters of the ANNs. The encoding depends on the parameters of the ANNs to be optimized (topology and/or weights).[43,44] The workflow of such a genetic algorithm is shown schematically in Figure 3. It starts with randomly initializing a population of individuals (Figure 3(a)), where each individual is evaluated and assigned a fitness value (Figure 3(b)) to measure its performance. The main part of the algorithm performs selection, crossover, and mutation operations aimed at iteratively improving the fitness values. The selection operator (Figure 3(c)) selects the individuals with better fitness values to construct a next generation of individuals. One of the most commonly used selection operators is roulette wheel selection, in which individuals are selected with a probability proportional to their fitness values. The stochasticity of this selection process may occasionally cause the best individuals to disappear from the population. It can be combined with the elitist selection scheme, which selects the top l best ranked individuals, such as n_1 and n_2 in Figure 3(c), and transfers them unchanged directly to the next generation. The crossover operator combines two individuals selected by the roulette wheel operator (parents, n_3 and n_4 in Figure 3(d)), to generate two new individuals (offspring, n' and n''). In particular, the 1-point crossover operator selects, with a probability of p_c, a random point to copy two different parts of two parents to generate offspring. Subsequently, a mutation operator flips the bit value in components of the binary representation of the offspring individuals randomly with a small probability p_m. Overall, the GA is run for a certain number of iterations or until a satisfactory solution, defined, e.g., by a threshold fitness value, is found.

3.3. Construction of a Deep FFNN for Prediction of Electronic Coupling Elements. In this work, we use a genetic algorithm to optimize the topology parameters (the number of hidden layers and the number of neurons in each hidden layer) of the feedforward neural networks. Each individual $n_i \in \mathbb{K}$ in the population is represented as five dimensional strings, where

$$\mathbb{K} = \{n := (i_1, \ldots, i_5) | i_j \in \mathcal{N} := \{0,50,100, \ldots, 1000\}\} \tag{8}$$

DOI: 10.1021/acs.jctc.8b01285
J. Chem. Theory Comput. 2019, 15, 1777−1784

which encodes an ANN topology. Therefore, the maximum number of hidden layers a network can have is set to 5, and the number of neurons each hidden layer is taken from the set of 21 discrete values. We limit the FFNN topologies in this manner to reduce the search space and computational complexity. Consequently, our genetic algorithm aims to find the optimum model settings in 21^5 = 4084101 number of possible networks.

The FFNNs are trained on a training data set using backpropagation to minimize the error between the target $\log_{10}[(J_{ij}^{\mathrm{DFT}}/\mathrm{eV})^2]$ (actual labels of the input data) and predicted outputs $\log_{10}[(J_{ij}^{\mathrm{FFNN}}/\mathrm{eV})^2]$. We use three distinct snapshots extracted from different MD simulations for training and validation. Each snapshot contains 4096 Alq_3 molecules with approximately 24000 pairs in the neighbor list as described in section 2. The first snapshot is used to optimize the weights of a chosen neural network, while the second snapshot is used for selecting the neural network architectures based on their fitness values. The third data set is used to analyze the performance of the constructed final neural network on completely unseen data.

The fitness value of a given feedforward neural network architecture is determined by evaluating the mean squared error of the difference between the mobility μ^{FFNN} and the reference mobility μ^{DFT}

$$\Xi = \left(\frac{1}{N_{\mathbf{F}}} \sum_{k=1}^{N_{\mathbf{F}}} (\mu_k^{\mathrm{DFT}} - \mu_k^{\mathrm{FFNN}})^2 \right)^{-1} \tag{9}$$

where $N_{\mathbf{F}}$ stands for the number of field values. For each $|\mathbf{F}| = b \cdot 10^7$ V/m with $b \in \{3,4,5,7,9\}$, the mobility is obtained as an average over $\pm x$, $\pm y$, $\pm z$ directions of the applied field.

Our GA starts with randomly initializing 30 individuals and evaluating the fitness of the constructed FFNNs, respectively, as illustrated in Figure 3. We use the roulette wheel selection with the elitism of $l = 2$ and standard 1-point crossover operators with the probability of $p_c = 1$. The mutation operator selects a random component of a string with a probability p_m = 0.1 and replaces it with a randomly selected value from $\mathcal{N}$. The probability of selecting $0 \in \mathcal{N}$ for the mutation is 0.3, and the rest of the values share 0.7 with equal probabilities to encourage smaller networks. The GA was run until there is no fitness improvement for 20 generations.

4. RESULTS

In Figure 4 we show the calculated field-dependence of the hole mobility in Alq_3 for various different models (a) without ($\mu(\Delta E_{ij} = 0)$) and (b) with ($\mu(\Delta E_{ij})$) energetic disorder taken into account in eq 2. The data indicated by the orange squares has been obtained by the ab initio model as described in section 2 and serves as the reference for the FFNN model.

In particular, we chose the disorder-free case as in Figure 4(a) in the evaluation of the fitness (eq 9) during the evolutionary FFNN optimization. Here, the rates and concomitantly the mobility are solely determined by the topological connectivity of the charge transporting network given by the electronic coupling elements.[45] The reference mobility has a minimally negative slope with increasing field strength, which is attributed to the system being in the inverted Marcus regime for $\Delta E_{ij} = 0$. Light green triangles show $\mu^{\mathrm{FFNN}}(F)$ as it results from two individuals in the first generation FFNN, with vastly different performances. The first

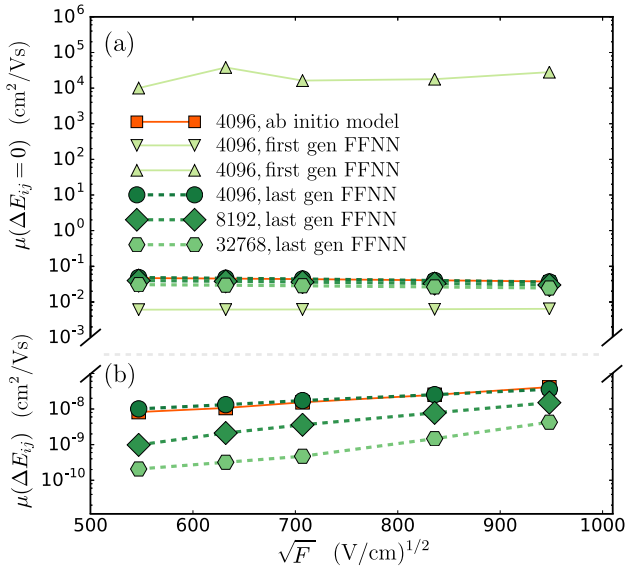


Figure 4. Field-dependent hole mobility (Poole-Frenkel plot) of Alq_3, for systems containing 4096, 8192, and 32768 molecules. In (a) the mobility μ for the disorder-free case, i.e. ΔE_{ij} = 0, is given, whereas (b) illustrates the mobility μ for the case with disorder, i.e., $\Delta E_{ij} \neq 0$.

one yields completely unphysical behavior with mobilities on the order of 10^4 cm^2/(V s), about 5 orders of magnitude larger than the ab initio reference. In comparison, the second model is much closer but underestimates $\mu^{\mathrm{DFT}}(F)$ consistently by about a factor of 10. While this agreement appears acceptable, a closer inspection of Figure 5(a) reveals a low fitness value (Ξ = 1.5 × 10^3 V^2 s^2/cm^4). The predicted $\log_{10}[(J^{\mathrm{FFNN}}/\mathrm{eV})^2]$ shows a MAE of 1.80 and is only weakly correlated to the DFT reference, as can be seen in Figure 5(c). From this starting point, the evolution of the FFNN results in an initially slowly increasing fitness. Going through 25 generations the fitness

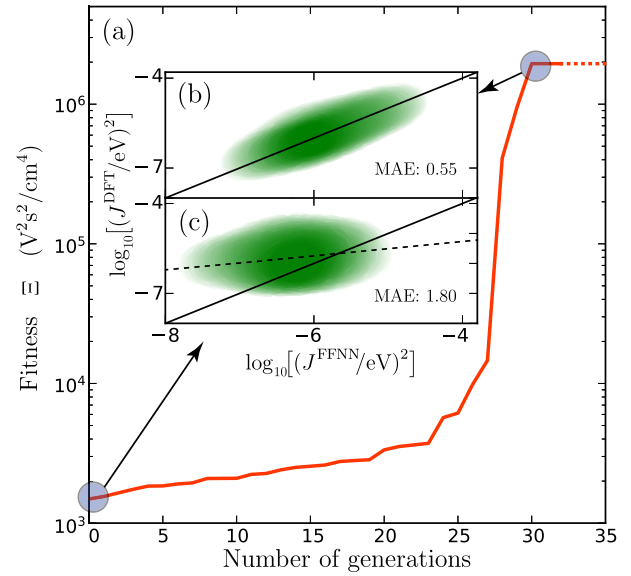


Figure 5. (a) Fitness evolution of the best performing feedforward neural network in each generation, showing a slow growth followed by rapid improvement reminiscent of punctuated equilibrium. (b) Correlation of predicted and reference data for the coupling elements of the final optimal FFNN model, compared to (c) the one in the first generation.

DOI: 10.1021/acs.jctc.8b01285
J. Chem. Theory Comput. 2019, 15, 1777−1784

only improves by a factor of 2. This slow growth is followed by a rapid fitness evolution that takes place within only six generations, after which it stops quickly, and the process ends up in an equilibrium situation. Such a phenomenon of instantaneous change is not unique to our evolutionary FFNN, and it has also been observed in evolutionary biology with similar patterns in the fossil records known as punctuated equilibrium.[46] The last generation FFNN consists of two layers with 800 and 550 neurons, respectively, and is characterized by a fitness value of $\Xi = 2 \times 10^6$ V^2 s^2/cm^4, an improvement of 3 orders of magnitude over the first model. This is also reflected by the obtained hole mobility (circles) in Figure 4(a), which is practically indistinguishable from the ab initio reference, see also Table 1. The MAE is reduced to 0.55, and the correlation in Figure 5(b) is clearly improved.

With a FFNN model at hand that shows good characteristics and performs well for the disorder-free case, we now evaluate its applicability in the realistic scenario with energetic disorder taken into account. This constitutes an important test for the FFNN model of the electronic coupling elements. While optimization of the model based on the disorder-free case should, ideally, predict the connectivity of the charge transporting network accurately, it does so for a completely flat energy landscape. It cannot be excluded a priori that coupling elements that are below the percolation threshold ($J^2 < 4 \times 10^{-7}$ eV^2), and hence insignificant for $\mu(\Delta E_{ij} = 0)$, provide low-probability, but crucial, additional pathways to avoid or escape low-energy regions (traps) in the $\Delta E_{ij} \neq 0$ case. With the large energetic disorder of $\sigma = 0.20$ eV the ab initio reference model exhibits a mobility reduction of about 6 orders of magnitude, see orange squares in Figure 4(b), and a noticeable positive field-dependence known as Poole-Frenkel behavior with $\mu(F) = \mu_0 \exp(\alpha F)$. The FFNN model reproduces this behavior extremely well, with an observed maximum error of 5.0×10^{-9} cm^2/(V s) and a MAE of 2.7×10^{-9} cm^2/(V s), which are both smaller than the average standard error of the mean of μ^{DFT} of 5.4×10^{-9} cm^2/(V s), see Table 1. Poole-Frenkel slopes of $\alpha^{\mathrm{DFT}} = 4.2 \times 10^{-3}$ (cm/V)$^{1/2}$ and $\alpha^{\mathrm{FFNN}} = 3.2 \times 10^{-3}$ (cm/V)$^{1/2}$ are in reasonable agreement with each other, taking the bounds of the respective errors of the individual mobility values as given in Table 1 into account.

Based on this comparison, we conclude that the FFNN provides very reliable predictions of electronic coupling elements in Alq_3 over a wide range of magnitudes taking into account explicit details of molecular orientations in large-scale morphologies. It also comes at a significantly reduced computational cost compared to the ab initio model. For a single frame containing N molecules and (on average) κN hopping pairs, the total CPU time for the coupling elements is $T_N = \kappa N T_{\mathrm{coupl}}$, where T_{coupl} is a typical CPU time per coupling element. Using the DFT based method as described in section 2 consumes about $T_{\mathrm{coupl}} = 20$ min on one thread of an Intel(R) Xeon(R) CPU E7-4830 v4 @ 2.00 GHz for Alq_3. With $\kappa \approx 5.5$, one obtains $T_{4096} = 312$ d. These calculations are however easily parallelizable in high-throughput settings, so that transport simulations can be performed within an acceptable total simulation time of, e.g., 1 week. For the 4096 molecule system of Alq_3, this can be achieved by using about 45 threads simultaneously. It is apparent that due to the linear scaling of T_N with the number of molecules, studying even slightly larger systems (which might be necessary when transport properties are system-size dependent or to investigate realistic carrier

Table 1. Hole Mobilities (in cm^2/(V s)) of Alq_3 for Different Values of Externally Applied Electric Field[a]

		3×10^{-7} V/m	4×10^{-7} V/m	5×10^{-7} V/m	7×10^{-7} V/m	9×10^{-7} V/m
			without disorder $\Delta E_{ij} = 0$			
4096	DFT	$4.7 \times 10^{-2} \pm 2.1 \times 10^{-4}$	$4.6 \times 10^{-2} \pm 1.7 \times 10^{-4}$	$4.4 \times 10^{-2} \pm 2.0 \times 10^{-4}$	$4.2 \times 10^{-2} \pm 2.1 \times 10^{-4}$	$4.1 \times 10^{-2} \pm 4.3 \times 10^{-4}$
4096	FFNN	$4.6 \times 10^{-2} \pm 1.8 \times 10^{-4}$	$4.5 \times 10^{-2} \pm 2.1 \times 10^{-4}$	$4.3 \times 10^{-2} \pm 2.5 \times 10^{-4}$	$4.0 \times 10^{-2} \pm 3.0 \times 10^{-4}$	$3.7 \times 10^{-2} \pm 4.3 \times 10^{-4}$
8192	FFNN	$3.9 \times 10^{-2} \pm 3.6 \times 10^{-4}$	$3.8 \times 10^{-2} \pm 3.8 \times 10^{-4}$	$3.6 \times 10^{-2} \pm 4.0 \times 10^{-4}$	$3.3 \times 10^{-2} \pm 4.3 \times 10^{-4}$	$3.0 \times 10^{-2} \pm 4.4 \times 10^{-4}$
32769	FFNN	$3.0 \times 10^{-2} \pm 5.1 \times 10^{-5}$	$2.9 \times 10^{-2} \pm 6.0 \times 10^{-5}$	$2.8 \times 10^{-2} \pm 6.7 \times 10^{-5}$	$2.6 \times 10^{-2} \pm 7.6 \times 10^{-5}$	$2.4 \times 10^{-2} \pm 7.8 \times 10^{-5}$
			with disorder $\Delta E_{ij} \neq 0$			
4096	DFT	$8.2 \times 10^{-9} \pm 1.5 \times 10^{-9}$	$1.1 \times 10^{-8} \pm 2.3 \times 10^{-9}$	$1.5 \times 10^{-8} \pm 4.4 \times 10^{-9}$	$3.0 \times 10^{-8} \pm 8.8 \times 10^{-9}$	$4.1 \times 10^{-8} \pm 9.8 \times 10^{-9}$
4096	FFNN	$1.0 \times 10^{-8} \pm 4.6 \times 10^{-10}$	$1.3 \times 10^{-8} \pm 1.2 \times 10^{-9}$	$1.7 \times 10^{-8} \pm 1.9 \times 10^{-9}$	$2.5 \times 10^{-8} \pm 2.6 \times 10^{-9}$	$3.6 \times 10^{-8} \pm 4.2 \times 10^{-9}$
8192	FFNN	$9.8 \times 10^{-10} \pm 4.6 \times 10^{-10}$	$2.1 \times 10^{-9} \pm 1.2 \times 10^{-9}$	$3.6 \times 10^{-9} \pm 1.9 \times 10^{-9}$	$7.8 \times 10^{-9} \pm 2.6 \times 10^{-9}$	$1.5 \times 10^{-8} \pm 4.2 \times 10^{-9}$
32769	FFNN	$2.1 \times 10^{-10} \pm 1.2 \times 10^{-10}$	$3.2 \times 10^{-10} \pm 1.7 \times 10^{-10}$	$4.7 \times 10^{-10} \pm 2.8 \times 10^{-10}$	$1.5 \times 10^{-9} \pm 1.1 \times 10^{-9}$	$4.3 \times 10^{-9} \pm 3.0 \times 10^{-9}$

[a]Results for cases both without ($\Delta E_{ij} = 0$) and with energetic disorder ($\Delta E_{ij} \neq 0$) are given for the three different system sizes considered, as obtained from DFT and FFNN based coupling elements, respectively.

DOI: 10.1021/acs.jctc.8b01285
J. Chem. Theory Comput. 2019, 15, 1777−1784

concentrations) implies a linear increase in either total simulation time or number of simultaneously used threads. Note that we are not addressing issues related to memory or storage. Compared to the cost of the DFT calculations of the reference data, the cost for training the FFNN is small. Within the search space, training of a single neural network, making use of a Nvidia Titan Xp GPU (3840 CUDA cores running at 1.6 GHz), takes about 5 min. Solving the Master equation (eq 4) for fitness evaluation as part of the EA approach was performed on a single CPU thread in a few seconds. The complete training procedure as in Figure 5 takes in total about 3 days. With the trained FFNN at hand, the evaluation of a coupling element is practically instantaneous, which removes the above costs and restrictions of the ab initio model.

To demonstrate the applicability of the FFNN in this context, we have also simulated Alq_3morphologies containing 8192 and 32769 molecules, respectively, following the same procedure as before, except for the calculation of the J_{ij}^{DFT}, which would have taken T_{8192} = 624 d and T_{32769} = 2496 d. Applying the FFNN model, the hole mobilities are calculated, and the results are shown in Figure 4 and in Table 1. In the disorder-free case (Figure 4(a)), the mobility is as expected practically independent of system size. With energetic disorder taken into account, the situation in Figure 4(b) is markedly different. Doubling the system size from 4096 to 8192 molecules lowers the mobility by about 1 order of magnitude, while another quadrupling further reduces the mobility by the same amount. Such a behavior is indicative of dispersive transport[47] and is related to the fact that the mean transport energy of the charge carrier depends on the system size.

All in all, the FFNN constructed with the evolutionary approach described in this work based on fitness evaluation in the $\Delta E_{ij} = 0$ case has not only proven to work well for the more realistic, unseen $\Delta E_{ij} \neq 0$ simulation but also in application to larger systems that are inaccessible to the standard ab initio model.

5. SUMMARY

To summarize, we have presented a general framework for the construction of a deep feedforward neural network to predict electronic coupling elements. The final FFNN model constructed for an amorphous organic semiconductor, tris(8-hydroxyquinoline)aluminum, showed good agreement with ab initio reference data with and without energetic disorder. Additionally, we have shown that the final model is applicable to larger systems, which makes the presented approach a promising candidate to overcome system size limitations inherent to computationally expensive multiscale approaches.

AUTHOR INFORMATION

Corresponding Authors

*E-mail: o.caylak@tue.nl.
*E-mail: a.yaman@tue.nl.
*E-mail: B.Baumeier@tue.nl.

ORCID

Björn Baumeier: 0000-0002-6077-0467

Funding

This work has been supported by the Innovational Research Incentives Scheme Vidi of The Netherlands Organisation for Scientific Research (NWO) with project number 723.016.002 and also by the European Union's Horizon 2020 research and innovation program under grant agreement No: 665347. We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan Xp Pascal GPU used for this research.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Jens Wehner and Gianluca Tirimbò for a critical reading of the manuscript.

REFERENCES

(1) Brédas, J.-L.; Beljonne, D.; Coropceanu, V.; Cornil, J. Charge-Transfer and Energy-Transfer Processes in ?-Conjugated Oligomers and Polymers: A Molecular Picture. *Chem. Rev.* **2004**, *104*, 4971−5004.

(2) Bakulin, A. A.; Rao, A.; Pavelyev, V. G.; van Loosdrecht, P. H. M.; Pshenichnikov, M. S.; Niedzialek, D.; Cornil, J.; Beljonne, D.; Friend, R. H. The Role of Driving Energy and Delocalized States for Charge Separation in Organic Semiconductors. *Science* **2012**, *335*, 1340−1344.

(3) Poelking, C.; Tietze, M.; Elschner, C.; Olthof, S.; Hertel, D.; Baumeier, B.; Würthner, F.; Meerholz, K.; Leo, K.; Andrienko, D. Impact of Mesoscale Order on Open-Circuit Voltage in Organic Solar Cells. *Nat. Mater.* **2015**, *14*, 434−439.

(4) Duan, C.; Willems, R. E. M.; van Franeker, J. J.; Bruijnaers, B. J.; Wienk, M. M.; Janssen, R. A. J. Effect of Side Chain Length on the Charge Transport, Morphology, and Photovoltaic Performance of Conjugated Polymers in Bulk Heterojunction Solar Cells. *J. Mater. Chem. A* **2016**, *4*, 1855−1866.

(5) Borsenberger, P. M.; Pautmeier, L.; Bässler, H. Charge Transport in Disordered Molecular Solids. *J. Chem. Phys.* **1991**, *94*, 5447.

(6) Bässler, H. Charge Transport in Disordered Organic Photoconductors a Monte Carlo Simulation Study. *Phys. Status Solidi B* **1993**, *175*, 15−56.

(7) Novikov, S. V.; Dunlap, D. H.; Kenkre, V. M.; Parris, P. E.; Vannikov, A. V. Essential Role of Correlations in Governing Charge Transport in Disordered Organic Materials. *Phys. Rev. Lett.* **1998**, *81*, 4472.

(8) Pasveer, W. F.; Cottaar, J.; Tanase, C.; Coehoorn, R.; Bobbert, P. A.; Blom, P. W. M.; de Leeuw, D. M.; Michels, M. A. J. Unified Description of Charge-Carrier Mobilities in Disordered Semiconducting Polymers. *Phys. Rev. Lett.* **2005**, *94*, 206601.

(9) Coehoorn, R.; Pasveer, W. F.; Bobbert, P. A.; Michels, M. A. J. Charge-Carrier Concentration Dependence of the Hopping Mobility in Organic Materials with Gaussian Disorder. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2005**, *72*, 155206.

(10) Cottaar, J.; Bobbert, P. A. Calculating Charge-Carrier Mobilities in Disordered Semiconducting Polymers: Mean Field and Beyond. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2006**, *74*, 115204.

(11) Rühle, V.; Lukyanov, A.; May, F.; Schrader, M.; Vehoff, T.; Kirkpatrick, J.; Baumeier, B.; Andrienko, D. Microscopic Simulations of Charge Transport in Disordered Organic Semiconductors. *J. Chem. Theory Comput.* **2011**, *7*, 3335−3345.

(12) Beljonne, D.; Cornil, J.; Muccioli, L.; Zannoni, C.; Brédas, J.-L.; Castet, F. Electronic Processes at Organic-Organic Interfaces: Insight from Modeling and Implications for Opto-Electronic Devices†. *Chem. Mater.* **2011**, *23*, 591−609.

(13) de Vries, X.; Friederich, P.; Wenzel, W.; Coehoorn, R.; Bobbert, P. A. Full Quantum Treatment of Charge Dynamics in Amorphous Molecular Semiconductors. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2018**, *97*, 075203.

(14) Risko, C.; McGehee, M. D.; Brédas, J.-L. A Quantum-Chemical Perspective into Low Optical-Gap Polymers for Highly-Efficient Organic Solar Cells. *Chemical Science* **2011**, *2*, 1200−1218.

(15) Schrader, M.; Fitzner, R.; Hein, M.; Elschner, C.; Baumeier, B.; Leo, K.; Riede, M.; Baeuerle, P.; Andrienko, D. Comparative Study of

DOI: 10.1021/acs.jctc.8b01285
J. Chem. Theory Comput. 2019, 15, 1777−1784

Microscopic Charge Dynamics in Crystalline Acceptor-Substituted Oligothiophenes. *J. Am. Chem. Soc.* **2012**, *134*, 6052−6056.
(16) May, F.; Baumeier, B.; Lennartz, C.; Andrienko, D. Can Lattice Models Predict the Density of States of Amorphous Organic Semiconductors? *Phys. Rev. Lett.* **2012**, *109*, 136401.
(17) Valeev, E. F.; Coropceanu, V.; da Silva Filho, D. A.; Salman, S.; Brédas, J.-L. Effect of Electronic Polarization on Charge-Transport Parameters in Molecular Organic Semiconductors. *J. Am. Chem. Soc.* **2006**, *128*, 9882−9886.
(18) Baumeier, B.; Kirkpatrick, J.; Andrienko, D. Density-Functional Based Determination of Intermolecular Charge Transfer Properties for Large-Scale Morphologies. *Phys. Chem. Chem. Phys.* **2010**, *12*, 11103−11113.
(19) Baumeier, B.; Stenzel, O.; Poelking, C.; Andrienko, D.; Schmidt, V. Stochastic Modeling of Molecular Charge Transport Networks. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2012**, *86*, 184202.
(20) Kordt, P.; Stenzel, O.; Baumeier, B.; Schmidt, V.; Andrienko, D. Parametrization of Extended Gaussian Disorder Models from Microscopic Charge Transport Simulations. *J. Chem. Theory Comput.* **2014**, *10*, 2508−2513.
(21) Schütt, O.; VandeVondele, J. Machine Learning Adaptive Basis Sets for Efficient Large Scale Density Functional Theory Simulation. *J. Chem. Theory Comput.* **2018**, *14*, 4168−4175.
(22) Misra, M.; Andrienko, D.; Baumeier, B.; Faulon, J.-L.; von Lilienfeld, O. A. Toward Quantitative Structure−Property Relationships for Charge Transfer Rates of Polycyclic Aromatic Hydrocarbons. *J. Chem. Theory Comput.* **2011**, *7*, 2549−2555.
(23) Hansen, K.; Montavon, G.; Biegler, F.; Fazli, S.; Rupp, M.; Scheffler, M.; von Lilienfeld, O. A.; Tkatchenko, A.; Müller, K.-R. Assessment and Validation of Machine Learning Methods for Predicting Molecular Atomization Energies. *J. Chem. Theory Comput.* **2013**, *9*, 3404−3419.
(24) Ramakrishnan, R.; Dral, P. O.; Rupp, M.; von Lilienfeld, O. A. Quantum Chemistry Structures and Properties of 134 Kilo Molecules. *Sci. Data* **2014**, *1*, 140022.
(25) Ramakrishnan, R.; Hartmann, M.; Tapavicza, E.; von Lilienfeld, O. A. Electronic Spectra from TDDFT and Machine Learning in Chemical Space. *J. Chem. Phys.* **2015**, *143*, 084111.
(26) Rupp, M.; Tkatchenko, A.; Müller, K.-R.; von Lilienfeld, O. A. Fast and Accurate Modeling of Molecular Atomization Energies with Machine Learning. *Phys. Rev. Lett.* **2012**, *108*, 058301.
(27) Haykin, S. O. *Neural Networks and Learning Machines*; Pearson Education: 2011.
(28) Goldberg, D. E. *Genetic Algorithms in Search, Optimization and Machine Learning*, 1st ed.; Addison-Wesley Longman Publishing Co., Inc.: Boston, MA, USA, 1989.
(29) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The Missing Term in Effective Pair Potentials. *J. Phys. Chem.* **1987**, *91*, 6269−6271.
(30) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.
(31) Lukyanov, A.; Lennartz, C.; Andrienko, D. Amorphous Films of Tris(8-Hydroxyquinolinato)Aluminium: Force-Field, Morphology, and Charge Transport. *Phys. Status Solidi A* **2009**, *206*, 2737−2742.
(32) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, Flexible, and Free. *J. Comput. Chem.* **2005**, *26*, 1701−1718.
(33) Marcus, R. A. Electron Transfer Reactions in Chemistry. Theory and Experiment. *Rev. Mod. Phys.* **1993**, *65*, 599.
(34) Becke, A. D. Density-functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(35) Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297.
(36) Thole, B. Molecular Polarizabilities Calculated with a Modified Dipole Interaction. *Chem. Phys.* **1981**, *59*, 341−350.
(37) van Duijnen, P. T.; Swart, M. Molecular and Atomic Polarizabilities: Thole's Model Revisited. *J. Phys. Chem. A* **1998**, *102*, 2399−2407.
(38) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868.
(39) Neese, F. The ORCA Program System. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2012**, *2*, 73−78.
(40) Wehner, J.; Brombacher, L.; Brown, J.; Junghans, C.; Çaylak, O.; Khalak, Y.; Madhikar, P.; Tirimbò, G.; Baumeier, B. Electronic Excitations in Complex Molecular Environments: Many-Body Green's Functions Theory in VOTCA-XTP. *J. Chem. Theory Comput.* **2018**, *14*, 6253−6268.
(41) Elton, D. C.; Boukouvalas, Z.; Butrico, M. S.; Fuge, M. D.; Chung, P. W. Applying Machine Learning Techniques to Predict the Properties of Energetic Materials. *Sci. Rep.* **2018**, *8*, 9059.
(42) Floreano, D.; Dürr, P.; Mattiussi, C. Neuroevolution: From Architectures to Learning. *Evolutionary Intelligence* **2008**, *1*, 47−62.
(43) Yaman, A.; Mocanu, D. C.; Iacca, G.; Fletcher, G.; Pechenizkiy, M. Limited Evaluation Cooperative Co-Evolutionary Differential Evolution for Large-Scale Neuroevolution. *Proceedings of the Genetic and Evolutionary Computation Conference*; New York, NY, USA, 2018; pp 569−576, DOI: 10.1145/3205455.3205555.
(44) Miikkulainen, R.; Liang, J.; Meyerson, E.; Rawal, A.; Fink, D.; Francon, O.; Raju, B.; Shahrzad, H.; Navruzyan, A.; Duffy, N.; Hodjat, B. Evolving Deep Neural Networks. In *Artificial Intelligence in the Age of Neural Networks and Brain Computing*; Kozma, R., Alippi, C., Choe, Y., Morabito, F. C., Eds.; Academic Press: 2019; pp 293−312, ISBN 978-0-12-815480-9, DOI: 10.1016/B978-0-12-815480-9.00015-3.
(45) Vehoff, T.; Baumeier, B.; Troisi, A.; Andrienko, D. Charge Transport in Organic Crystals: Role of Disorder and Topological Connectivity. *J. Am. Chem. Soc.* **2010**, *132*, 11702−11708.
(46) Ayala, F. J.; Avise, J. C. *Essential Readings in Evolutionary Biology*; JHU Press: 2014.
(47) Borsenberger, P.; Pautmeier, L.; Bässler, H. Nondispersive-to-Dispersive Charge-Transport Transition in Disordered Molecular Solids. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1992**, *46*, 12145−12153.

DOI: 10.1021/acs.jctc.8b01285
J. Chem. Theory Comput. 2019, 15, 1777−1784