Connecting exciton diffusion with surface roughness via deep learning

Liyao Lyu,[†,‡] Zhiwen Zhang,[*,¶] and Jingrun Chen[*,†,§]

†*School of Mathematical Sciences, Soochow University, Suzhou, China*

‡*CW Chu College, Soochow University, Suzhou, China*

¶*Department of Mathematics, The University of Hong Kong, Pokfulam Road, Hong Kong SAR, China*

§*Mathematical Center for Interdisciplinary Research, Soochow University, Suzhou, China*

E-mail: zhangzw@hku.hk; jingrunchen@suda.edu.cn

Abstract

Exciton diffusion plays a vital role in the function of many organic semiconducting opto-electronic devices, where an accurate description requires precise control of heterojunctions. This poses a challenging problem because the parameterization of heterojunctions in high-dimensional random space is far beyond the capability of classical simulation tools. Here, we develop a novel method based on deep neural network to extract a function for exciton diffusion length on surface roughness with high accuracy and unprecedented efficiency, yielding an abundance of information over the entire parameter space. Our method provides a new strategy to analyze the impact of interfacial ordering on exciton diffusion and is expected to assist experimental design with tailored opto-electronic functionalities.

INTRODUCTION

Over the past decades, much attention has been paid on organic semiconductors for applications in various opto-electronic devices.[1–4] These materials include small molecules,[5,6] oligomers,[7,8] and polymers.[9,10] Exciton diffusion is one of the key processes behind the operation of organic opto-electronic devices.[11–13] From a microscopic perspective, exciton, a bound electron-hole pair, is the elementary excitation in opto-electronic devices such as light emitting diodes and organic solar cells. The exciton diffusion length (EDL) is the characteristic distance that excitons are able to travel during their lifetime.[6] A short diffusion length in organic photovoltaics limits the dissociation of excitons into free charge.[14,15] Conversely, a large diffusion length in organic light emitting diodes may limit luminous efficiency if excitons diffuse to non-radiative quenching sites.[16]

As quasi-particles with no net charge, excitons are difficult to probe directly by electrical means.[17] This is particularly true in organic semiconductors where the exciton binding energy is $\sim$1 electronvolt.[18] Reported techniques to measure EDL include photoluminescence (PL) surface quenching,[6,14,19–25] time-resolved PL bulk quenching modeled with a Monte Carlo simulation,[10,26] exciton-exciton annihilation,[27–30] modeling of solar cell photocurrent spectrum,[5,9,25,31–39] time-resolved microwave conductance,[40–42] spectrally resolved PL quenching[43–45] and Förster resonance energy transfer theory.[43,46,47] From a theoretical perspective, the minimal modeling error is given by the diffusion equation model,[48] which is employed in the current work.

To be precise, the device used in PL surface quenching experiment includes two layers of organic materials with thickness ranging from dozens of nanometers to hundreds of nanometers. One layer of material is called donor and the other is called acceptor or quencher according to the difference of their chemical properties. Under the illumination of solar lights, excitons are generated in the donor layer and diffuse in the donor. Due to the exciton-environment interaction, some excitions die out and emit photons which lead to the PL. The donor-acceptor interface serves as the absorbing boundary while other boundaries serve as

reflecting boundaries due to the tailored properties. Since the donor-acceptor interface is not exposed to the air/vacuum and the resolution of the surface morphology is limited by the resolution of atomic force microscopy, the interface is subject to an uncertainty. It is found that the fitted EDL is sensitive to the uncertainty in some scenarios. From a numerical perspective, the random interface requires a parametrization in high-dimensional random space, which is prohibitively expensive for any simulation tool. For example, Monte Carlo method overcomes the curse of dimensionality but has very low accuracy.[49] Stochastic collection method has high accuracy but is only affordable in low dimensional random space.[50] Asymptotics-based method is efficient but its accuracy relies heavily on the magnitude of randomness.[51] In the current work, we propose a novel method based on deep learning with high accuracy and unprecedented efficiency.

Recently, increasing attentions have been paid to apply machine learning (ML) techniques to materials-related problems. For example, the classification of crystal structures of transition metal phosphide via support vector machine[52] leads to the discovery of a novel phase.[53] Likewise, a hybrid probabilistic model based on high-throughput first-principle computation and ML was developed to identify stable novel compositions and their crystal structures.[54] Physical parameters such as band gap,[55,56] elastic constants,[55,57] and Debye temperature[55] have also been predicted using an array of ML techniques. In another line, deep learning (DL) in computer science has had great success in text classification,[58] computer vision,[59] natural language processing,[60] and other data-driven applications. One significant advantage of DL is its strong ability to approximate a complex function in high dimensions and extract features with high precision using composition of simple nonlinear units. Meanwhile, benefiting from recent advances in parallel graphics processing unit - accelerated computing, huge volumes of data can be put into the DL architecture for training.

In this work, we employ DL to extract a complex function of EDL in terms of the random interface parametrized in a high-dimensional space. The fitted function has rich information, which explains a few interesting experimental observations. Compared to classical simula-

tion tools, our approach has the following features: quasi-Monte Carlo sampling[61] for data collection and ResNet[62] for training. The size of data in the former step grows only linearly with respect to the dimension of random space, thus our approach overcomes the curse of dimensionality. With the usage of ResNet in the latter step, a complex function can be extracted with high accuracy. Therefore, results provided here are completely out of the capability of classical simulation tools.

METHODS

Our approach consists of four major components: quasi-Monte Carlo sampling over the high-dimensional random space; diffusion equation model for data generation; ResNet for training to approximate a complex function of EDL; Information extraction for analysis (Figure 1).

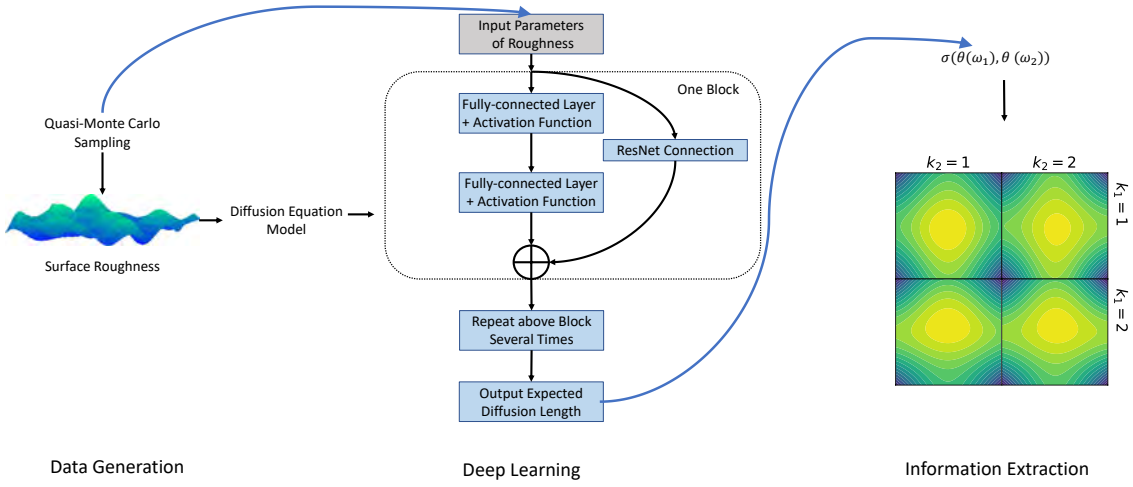


Figure 1: Flow chart of the deep learning method for extracting exciton diffusion length over the parameter space. Left: data generation; Middle: data training; Right: data prediction. In the stage of data generation, quasi-Monte Carlo method is used to sample the random space, and the actual exciton diffusion length is generated by solving the diffusion equation model. In the stage of data training, a complex function $\sigma(\theta(\omega_1), \theta(\omega_2))$ is approximated over the entire parameter space. In the stage of data prediction, given the full landscape of $\sigma(\theta(\omega_1), \theta(\omega_2))$, both qualitative and quantitative information can be extracted.

Model Description

An exciton that diffuses in the donor layer follows a diffusion-type equation over a 3D random domain $D = \{(x, y, z) \,|h(y, z, \omega_1, \omega_2) < x < d, 0 < y < L_y, 0 < z < L_z\}$. Here the donor-acceptor interface $x = h(y, z, \omega_1, \omega_2)$ is parameterized by

$$h(y, z, \omega_1, \omega_2) = \hat{h} \sum_{k_1=1}^{K_1} \sum_{k_2=1}^{K_2} k_1^{\beta} k_2^{\beta} \theta_{k_1}(\omega_1) \theta_{k_2}(\omega_2) \phi_{k_1}(y) \phi_{k_2}(z), \tag{1}$$

where $\hat{h}$ is the magnitude of length due to the roughness limited by the resolution of atomic force microscopy, $\theta_{k_1}(\omega_1), \theta_{k_2}(\omega_2)$ are i.i.d. random variables, $\phi_{k_1}(y) = \sin(2k_1\pi\frac{y}{L_y})$, $\phi_{k_2}(z) = \sin(2k_2\pi\frac{z}{L_z})$, and $\beta < 0$ controls the decay rate of spatial modes $\phi_{k_1}(y), \phi_{k_2}(z)$. The rougher the interface is, the closer the β approaches 0. Given the surface roughness measured in an experiment, parameters in (1) can be extracted via discrete Fourier transform.

The 3D diffusion equation reads as

$$\begin{cases} \sigma^2 \triangle u - u + G(x, y, z) = 0 & x \in D, \\ u_x(d, y, z) = 0 & 0 < y < L_y, 0 < z < L_z, \\ u(h(y, z, \omega_1, \omega_2), y, z) = 0 & 0 < y < L_y, 0 < z < L_z, \\ u(x, y, z) = u(x, y + L_y, z) = u(x, y, z + L_z) & h(y, z, \omega_1, \omega_2) < x < d, \end{cases} \tag{2}$$

where σ is the EDL to be extracted, u is the exciton density, G is the normalized exciton generation function by the transfer matrix method.[63] $x = d$ serves as the reflecting boundary and Neumann boundary condition is imposed on the boundary exposed in air. $x = h(y, z, \omega_1, \omega_2)$ serves as the absorbing boundary and homogenous Dirichlet boundary condition is imposed on the donor-acceptor interface. Periodic boundary conditions are imposed on in-plane directions y and z. For comparison and completeness, 1D and 2D models are given in the Supporting Information.

The PL is computed by

$$I_{\theta(\omega_1),\theta(\omega_2)}[\sigma, d] = \frac{1}{L_z}\frac{1}{L_y}\int_0^{L_z}\int_0^{L_y}\int_{h(y,z,\omega_1,\omega_2)}^{d} u(x,,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z. \tag{3}$$

For comparison with the 1D model, we divide the PL in the usual sense by lengths in x and y directions. In the experiment, PL data $\{\hat{I}_i\}_{i=1}^N$ are measured by a series of bilayer devices with different thicknesses $\{d_i\}_{i=1}^N$, where d_i is the thickness of the i-th donor layer.

The optimal EDL σ is expected to reproduce the experimental date $\{d_i, \hat{I}_i\}_{i=1}^N$ in the sense of minimized mean square error (MSE)

$$\min_{\sigma} J_{\theta(\omega_1),\theta(\omega_2)}(\sigma) = \frac{1}{N}\sum_{i=1}^{N}\left(I_{\theta(\omega_1),\theta(\omega_2)}(\sigma, d_i) - \hat{I}_i\right)^2. \tag{4}$$

Newton's method is used to solve (4) for σ (see the Supporting Information). The calculated σ is defined as $\sigma_{\theta(\omega_1),\theta(\omega_2)}$. Therefore, for different parameters $\theta(\omega_1), \theta(\omega_2)$, we get a data set $\left\{\left(\theta(\omega_1)[j], \theta(\omega_2)[j], \sigma_{\theta(\omega_1)[j],\theta(\omega_2)[j]}\right)_{j=1}^M\right\}$ with M the size of data set.

ResNet

ResNet[62] is used to approximate $\sigma_{\theta(\omega_1),\theta(\omega_2)}$. A ResNet consists of a series of blocks. One block is given in Figure 1 with two linear transformations, two activation functions, and one short cut. Detailed description of ResNet is included in the Supporting Information. Parameters of the surface roughness ($\theta(\omega_1)$,$\theta(\omega_2)$) are fed as input, and the EDL σ is extracted as the output function over the entire parameter space. Sigmoid function is chosen as the activation function here.

The loss function we use is the MSE between the actual EDL $\sigma_{\theta(\omega_1),\theta(\omega_2)}$ given by the diffusion equation model and the predicted EDL $\sigma(\theta(\omega_1)[j], \theta(\omega_2)[j])$ given by the ResNet

$$MSE = \frac{1}{M}\sum_{j=1}^{M}\left(\sigma_{\theta(\omega_1)[j],\theta(\omega_2)[j]} - \sigma(\theta(\omega_1)[j], \theta(\omega_2)[j])\right)^2, \tag{5}$$

where θ represents the parameter set in the ResNet, j is the j-th sample, and M is the size of training data set.

Define the relative L^∞ error of EDL as

$$Error = \max_{1\leq j\leq M} \frac{\left|\sigma_{\theta(\omega_1)[j],\theta(\omega_2)[j]} - \sigma(\theta(\omega_1)[j], \theta(\omega_2)[j])\right|}{\sigma_{\theta(\omega_1)[j],\theta(\omega_2)[j]}}, \tag{6}$$

which will be used to quantify the approximation accuracy of DL.

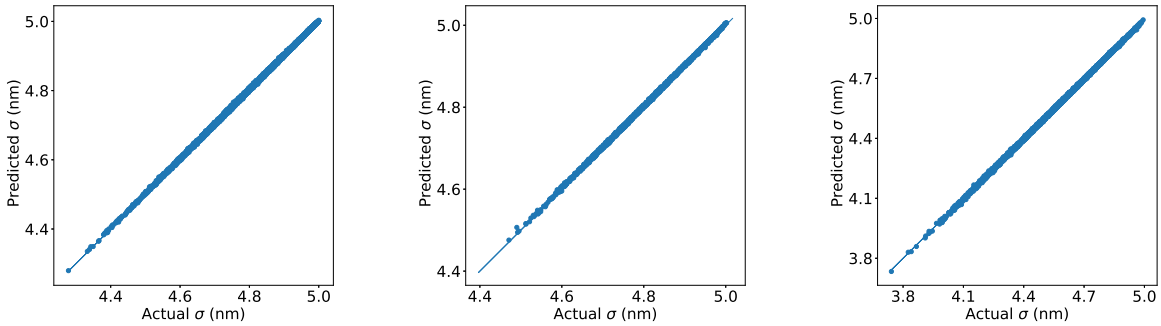
Quasi-Monte Carlo Sampling

Compared to uniform sampling and Monte-Carlo sampling, quasi-Monte Carlo sampling provides the best compromise between accuracy and efficiency. It overcomes the curse of dimensionality and has high accuracy.[61] For the simulations in our work, at least three orders of magnitude reduction in the size of data set is found for quasi-Monte Carlo sampling without loss of accuracy (see the Supporting Information).

Results and Discussion

Accuracy check and training data set

For the accuracy check in the 3D case, the reference PL data are generated using 5 realizations with out-of-plane thicknesses $d_i = 10, 15, 20, 25$ nm and $\sigma = 5$ nm in the absence of randomness. Afterwards, randomness is added with $K_1 = K_2 = 5$, i.e., $\theta(\omega_1)$ and $\theta(\omega_2)$ are arrays with 5 variables. Quasi-Monte Carlo sampling is used to generate 20000 points with the corresponding EDL obtained by solving (1) - (4). The first 15000 data are used as the training set, while the remaining data are used to check the predictability of the trained neural network; see Figure 2. Relative L^∞ errors of EDL are $0.270\%, 0.368\%$ and 0.532% for $\beta = -2, -1, 0$, respectively. It is known that the random field is closer to the white noise when $\beta = 0$ and thus is more difficult to be trained. However, uniform generalization errors



(a) $\beta = -2$ (b) $\beta = -1$ (c) $\beta = 0$

Figure 2: Uniform generalization error of the trained neural network for exciton diffusion length when $\beta = -2, -1, 0$. Relative L^∞ errors of exciton diffusion length are $0.270\%, 0.368\%$ and 0.532%, respectively.

for three different scenarios are observed, implying the robustness of trained neural network. Moreover, the size of training data set is small in the sense that only linear growth with respect to the dimension of random variables is observed, in contrast to other sampling techniques which either have the curse of dimensionality or low accuracy. Similar performance is observed for the 2D model (see the Supporting Information).

Information extraction

The trained neural network fits a high-dimensional function for EDL in terms of surface roughness. Rich information can be extracted based on the fitted function. We demonstrate this using three examples.

Modeling error Expectations of EDL in 3D are recorded in Table 1 for $\beta = -2, -1, 0$. PL data are generated using the 1D model with the reference EDL 5 nm. When $\beta = -2$, the EDL is close to 5 nm, which implies the equivalence between the 3D model and the 1D model. However, when $\beta = 0$, the EDL is clearly away from 5 nm. We attribute this difference to the modeling error between the 1D model and the 3D model with a surface roughness characterized by (1) with $\beta = 0$. So far, the 1D model is largely used in the literature to extract the EDL.[6,9,48,64] The main assumption underlying the modeling is the

high crystalline order of the organic material. When $\beta = -2$, long-range ordering exists in the random interface, which implicitly connects with the crystalline ordering of the material. Therefore, in this case, the 3D model and the 1D model are equivalent. However, when $\beta = 0$, only short-range ordering exists. As a consequence, the 3D model and the 1D model are not equivalent any more. Given a surface roughness from the experimental measurement, we can fit a function of form (1) using discrete Fourier transform, from which we can get the decay rate β and thus decide whether the 1D model is adequate or not. It is worth mentioning that similar results are observed in 2D using the asymptotics-based approach.[51]

Table 1: Expectations of exciton diffusion length in 3D for different surface roughness. The reference value is 5 nm.

$\beta = -2$	$\beta = -1$	$\beta = 0$
4.986 nm	4.842 nm	4.566 nm

Landscape exploration Contour plots of the fitted EDL on random variables are given in Figures 3 and 4 when $\beta = -2$, $\beta = 0$ and in Figure 17 when $\beta = -1$ (see the Supporting Information). In each subfigure, EDL σ is plotted as a function of $\theta_{k_1}(\omega_1)$ and $\theta_{k_2}(\omega_2)$, where $k_1, k_2 = 1, 2, 3, 4, 5$ and all the remaining random variables are set to be 0. A direct comparison between Figure 3 and Figure 4 illustrates the directional (anisotropic) dependence of EDL on random variables, due to different decay rates of random variables in the surface roughness.

Mode dependence Figure 5 provides a detailed demonstration of the dependence of EDL on random variables for $\beta = -2, -1, 0$. For illustration, we keep $\theta(\omega_2) = [1; 0; 0; 0; 0]$ fixed in the left column and $\theta(\omega_1) = [1; 0; 0; 0; 0]$ fixed in the right column. One distinct difference between 3D and 2D is that the maximum EDL is approached in the absence of randomness in 3D, in contrast to the minimum EDL in 2D (see the Supporting Information). The 3D result is reasonable since experimentally larger EDL is observed if the effect of surface roughness is minimized, while the 2D result is also of interest due to the unique

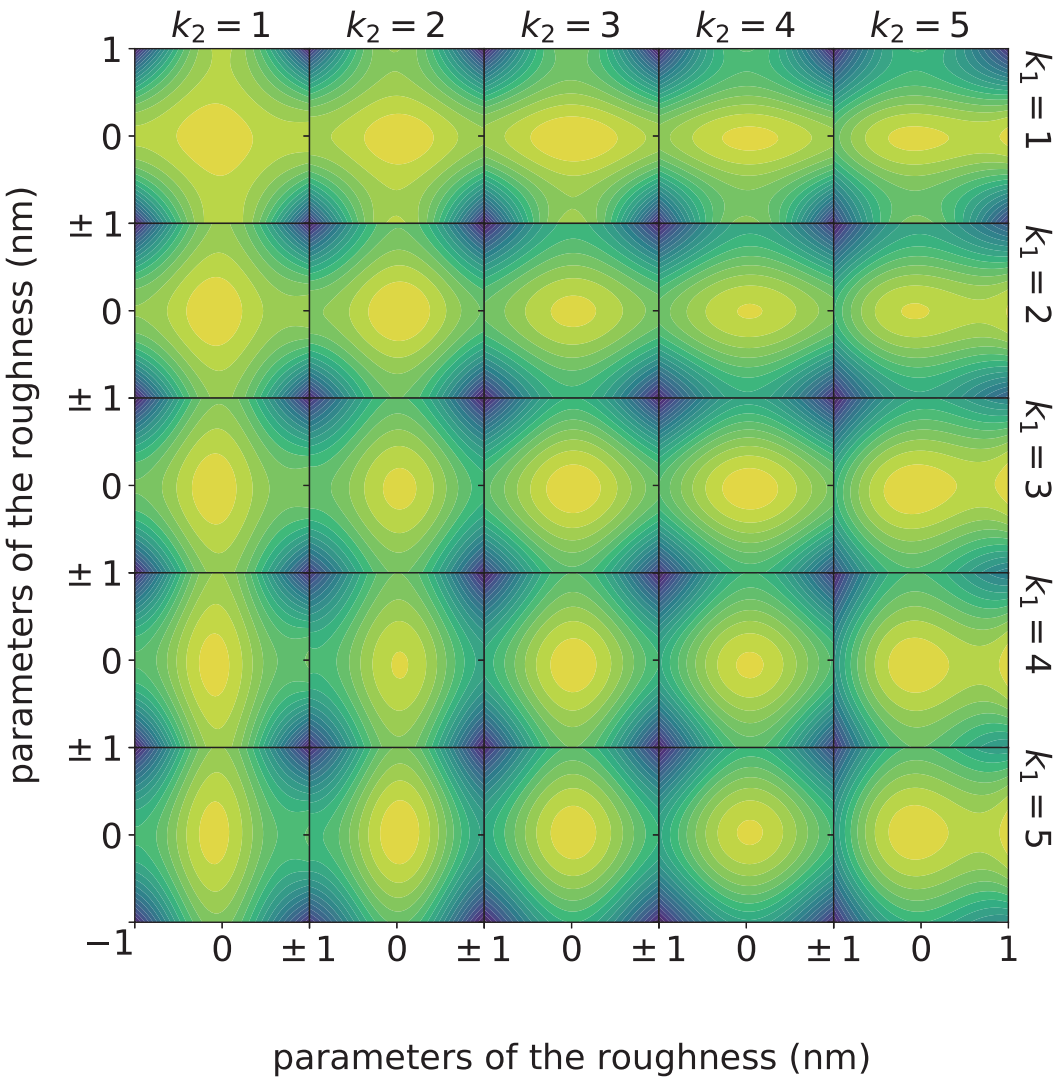


Figure 3: Contour plot of exciton diffusion length on random variables in 3D when $\beta = -2$.

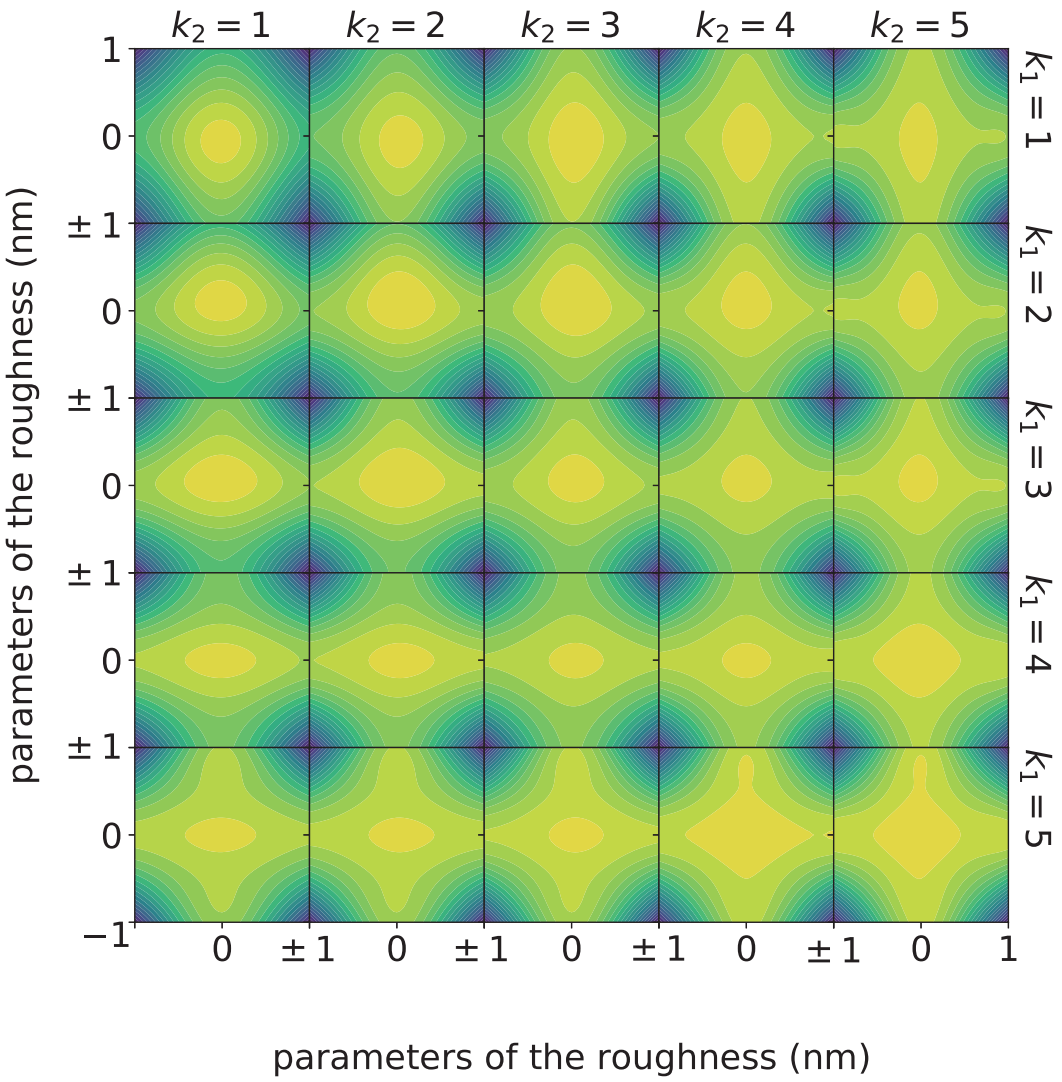


Figure 4: Contour plot of exciton diffusion length on random variables in 3D when $\beta = 0$.

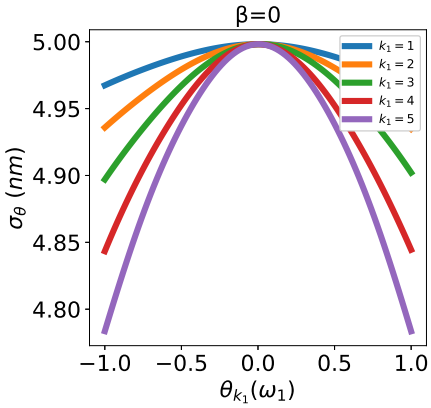
dimensional dependence. When $\beta = -2$, the EDL is more sensitive to the lower-order modes (smaller k) and is less sensitive to the high-order modes (larger k). When $\beta = 0$, the trend is completely opposite. This observation provides a detailed connection between surface roughness and EDL, which also sheds light on the experimental design. Given a surface roughness characterized by (1), we have the value of β, from which we know which mode is of the most importance. Consequently, targeted experimental techniques can be applied to improve the opto-electronic performance.

Conclusion

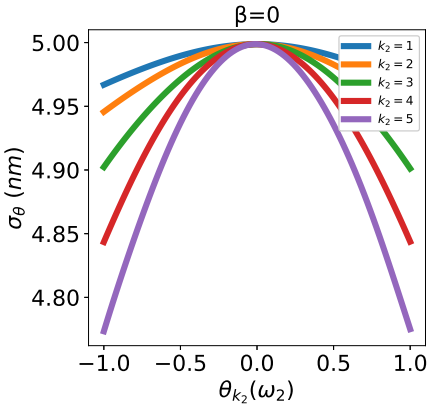
In summary, we have developed a novel method based on quasi-Monte Carlo sampling and ResNet to approximate the exciton diffusion length in terms of surface roughness parametrized by a high-dimensional random field. This method extracts a function for exciton diffusion length over the entire parameter space. Rich information, such as landscape profile and mode dependence, can be extracted with unprecedented details. Useful information regarding the modeling error and the experimental design can be provided, which sheds lights on how to reduce the modeling error and how to design better experiments to improve opto-electronic properties of organics materials.

Acknowledgments

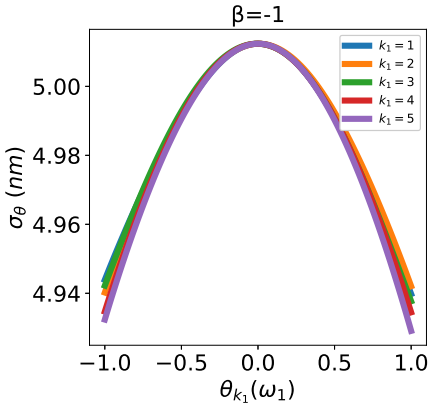
L. Lyu acknowledges the financial support of Undergraduate Training Program for Innovation and Entrepreneurship, Soochow University (Projection 201810285019Z). Z. Zhang acknowledges the financial support of Hong Kong RGC grants (Projects 27300616, 17300817, and 17300318) and National Natural Science Foundation of China via grant 11601457, Seed Funding Programme for Basic Research (HKU), and Basic Research Programme of The Science, Technology and Innovation Commission of Shenzhen Municipality (JCYJ20180307151603959). J. Chen acknowledges the financial support by National Natural Science Foundation of China



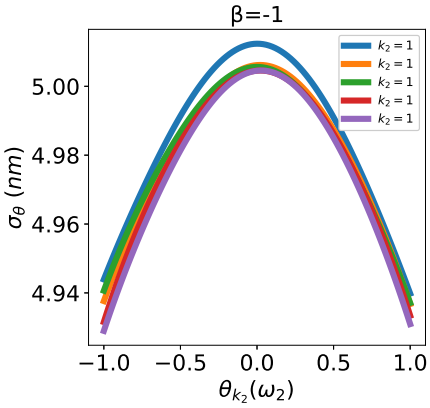
(a) $\theta(\omega_2)$ fixed



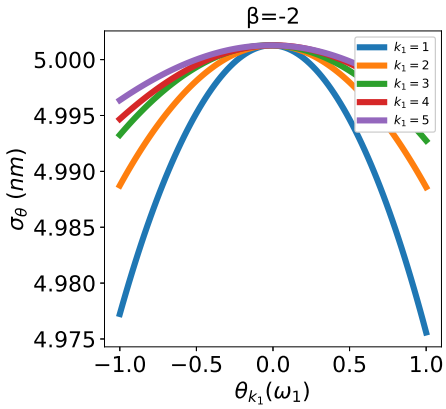
(b) $\theta(\omega_1)$ fixed



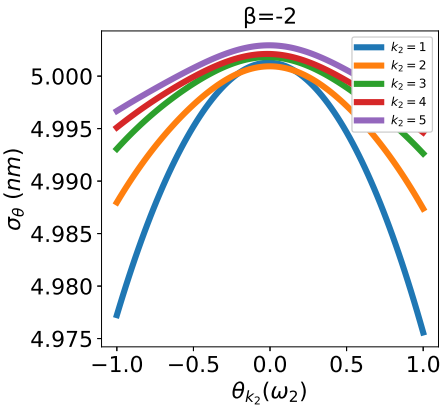
(c) $\theta(\omega_2)$ fixed



(d) $\theta(\omega_1)$ fixed



(e) $\theta(\omega_2)$ fixed



(f) $\theta(\omega_1)$ fixed

Figure 5: Detailed dependence of exciton diffusion length on random variables. Top row: $\beta = -2$; Middle row: $\beta = -1$; Bottom row: $\beta = 0$. Left column: $\theta(\omega_2) = [1; 0; 0; 0; 0]$ is fixed and σ is plotted as a function of $\theta_{k_1}(\omega_1) \in [-1, 1]$, where $k_1 = 1, 2, 3, 4, 5$; Right column: $\theta(\omega_1) = [1; 0; 0; 0; 0]$ is fixed and σ is plotted as a function of $\theta_{k_1}(\omega_1) \in [-1, 1]$, where $k_1 = 1, 2, 3, 4, 5$.

via grants 21602149 and 11971021, and the Science and Technology Development Fund, Macau SAR (File No. 101/2016/A3). Part of the work was done when J. Chen was visiting Department of Mathematics, University of Hong Kong. J. Chen would like to thank its hospitality.

Supporting Information Available

The following files are available free of charge.

Supporting Methods

1D and 2D models

The 1D model is defined over the domain

$$D_1 = \{x : x \in [\theta(\omega), d]\}, \tag{7}$$

where the interface is reduced to a random point $x = \theta(\omega)$.

The corresponding diffusion equation is

$$\begin{cases} \sigma^2 u_{xx} - u + G(x) = 0, \quad \theta(\omega) < x < d, \\ u_x(d) = 0, \\ u(\theta(\omega)) = 0, \end{cases} \tag{8}$$

and the photoluminescence (PL) is

$$I_{\theta(\omega)}[\sigma, d] = \int_{\theta(\omega)}^{d} u(x)\mathrm{d}x. \tag{9}$$

The 1D model is commonly used to extract the exciton diffusion length (EDL) due to its simplicity and model accuracy.[6,48]

The 2D model is defined over a random domain

$$D_2 = \{(x, y)\,|h(y, \omega) < x < d, 0 < y < L_y\}, \tag{10}$$

where the interface is a random line parametrized by $h(y, \omega) = \hat{h} \sum_{k=1}^{K} k^{\beta} \theta_k(\omega)\phi_k(y)$ with

$\phi_k(y) = \sin(\frac{2\pi ky}{L_y})$.

The corresponding diffusion equation is

$$\begin{cases} \sigma^2 \triangle u - u + G(x,y) = 0, & x \in D_2, \\ u_x(d) = 0, \quad u(h(y,\omega),y) = 0, & 0 < y < L_y, \\ u(x,y) = u(x,y+L_y), & h(y,\omega) < x < d, \end{cases} \tag{11}$$

and the PL is

$$I_{\theta(\omega)}[\sigma, d] = \frac{1}{L_y} \int_0^{L_y} \int_{h(y,\omega)}^{d} u(x,y) \mathrm{d}x \mathrm{d}y. \tag{12}$$

At the formal level, when $L_z \to 0$, the PL of 3D model defined by (3) reduces to the PL of 2D model defined by (12), and further they reduce to the PL of 1D model defined by (9) as $L_y \to 0$.

Newton's method

Given $\sigma^{(0)}$, for $n = 1, 2, \cdots$, until convergence, Newton's method for (4) solves

$$\sigma^{(n)} = \sigma^{(n-1)} - \alpha_n \frac{\frac{\partial}{\partial \sigma} J(\sigma^{(n-1)})}{\frac{\partial^2}{\partial^2 \sigma} J(\sigma^{(n-1)})}$$

with $\alpha_n \in (0, 1]$ given by line search.[65]

Given one realization of the random interface (1), by solving the 3D diffusion equation model (2) - (3), we get one datum $(\theta(\omega_1), \theta(\omega_2), \sigma_{\theta(\omega_1),\theta(\omega_2)})$, where $\theta(\omega_1)$ and $\theta(\omega_2)$ are inputs and $\sigma_{\theta(\omega_1),\theta(\omega_2)}$ is the output. A set of data $\left\{(\theta(\omega_1)[j], \theta(\omega_2)[j], \sigma_{\theta(\omega_1)[j],\theta(\omega_2)[j]})_{j=1}^{M}\right\}$ will be generated for training and testing.

Quasi-Monte Carlo sampling

In the sampling stage of data preparation, a large M is needed to ensure that the extracted function of EDL has the desired accuracy. There are two classical choices: uniform sampling

and random sampling. For uniform sampling, M grows exponentially fast with respect to K_1 and K_2. For example, in the 3D case, if $K_1 = K_2 = 5$ and points are uniformly distributed for each random variable, the size of training data set is shown in Table 2. Figure 6 plots

Table 2: Size of training data set for uniform sampling.

Number of points in each dimension	2	3	5	9
Size of training data set	1024	59049	9765625	3486784401

the points by uniform sampling when $K_1 = K_2 = 1$ (two random variables). Clearly such a sampling strategy has the curse of dimensionality.

On the other hand, if random sampling is used, then we do not have this issue. However, Monte-Carlo method has poor accuracy $\sim O(\frac{1}{\sqrt{M}})$. At least millions of data are needed for training. Meanwhile, for each datum, an inverse problem with the diffusion equation model over a curved domain in 3D has to be solved. These together make the network training prohibitively expensive. Fortunately, the quasi-Monte Carlo sampling has accuracy $\sim O(\frac{1}{M})$,[61] which reduces the size of training data set by orders of magnitudes in comparison with Monte-Carlo method. Specifically, we use Sobol sequence to generate points over the (high-dimensional) random space. Figure 7 plots the points generated by Sobol sequence, which is a deterministic way to generate points with better approximation accuracy. The size of data in the quasi-Monte Carlo method grows merely linearly fast with respect to the number of random variables. For the simulations in our work, at least three orders of magnitude reduction in the size of data set is found for quasi-Monte Carlo sampling strategy. Figures 8 and 9 show the huge advantage of quasi-Monte-Carlo sampling over uniform sampling. For the same size of training data set, the relative L^∞ error is 30.561% and 0.237%, implying more than two orders of magnitude improvement in the prediction accuracy.

Softwares

The following softwares and libraries are used: Julia, Flux and CuArrays. Julia is a high-level programming language designed for high-performance numerical analysis and computational science.[66] Flux is a library for machine learning. It comes "batteries-included" with many useful tools built in, but also allows taking the full power of the Julia language. CuArrays provides a fully-functional GPU array, which can give significant speedups over normal arrays without code changes.

Detailed description of ResNet

The ResNet network we use is stacked by several blocks with each block containing two linear transformations, two activation functions, and one shortcut connection. The i-th block can be expressed as

$$t = f_i(s) = g(W_{i,2} \cdot g(W_{i,1} \cdot s + b_{i,1}) + b_{i,2}) + s. \tag{13}$$

Here $s, t \in R^m$ are input and output of the i-th block, and weights $W_{i,j} \in R^{m \times m}, b_{i,1}, b_{i,2} \in R^m$. Sigmoid function

$$g(x) = \frac{1}{1 + \exp(-x)}$$

is chosen as the activation function to balance training complexity and accuracy.

The last term in (13) is called the shortcut connection or the residual connection. Advantages of using it are

1) It can solve the notorious problem of vanishing/exploding gradients automatically;

2) Without adding any parameters or computational complexity, the shortcut connection performing as an $Identity$ mapping can resolve the degradation issue (with the network depth increasing, accuracy gets saturated and then degrades rapidly).

The fully n-layer network can be expressed as

$$f_w(x) = f_n \circ f_{n-1} \cdots \circ f_1(x),$$

where w denotes the set of parameters in the whole network. Note that the input x in the first layer is in R^{dim} and the output of the whole structure $\sigma(\theta(\omega_1), \theta(\omega_2))$ is in R^1. To deal with the problem, we apply two linear transformations on both x before putting it into the ResNet structure and on the output of the ResNet structure. For example, we choose $m = 30, n = 6$ in the 3D model. Both $\theta(\omega_1)$ and $\theta(\omega_2)$ have 5 random variables, and thus dim = 10. Therefore, we apply two linear transforms: one from a 10 dimensional vector to a 30 dimensional vector and the other from a 30 dimensional vector to 1 dimensional vector before and after the ResNet structure. Parameters in these linear transforms also need to be trained.

2D results

First, we focus on the 2D problem with only one realization, i.e., only one $d = 10$ and $N = 1$. PL data are generated when $\sigma = 10$ without any randomness. Accuracy of the trained neural network in terms of size of the training set is recorded in Table 3. From the

Table 3: Generalization error of the trained neural network model for a random field with different decay rates in 2D.

Size of training data set	9	25	81
Error ($\beta = -2$)	1.638%	0.04466%	0.00209%
Error ($\beta = -1$)	0.101%	0.00795%	0.00234%
Error ($\beta = 0$)	0.765%	0.105%	0.0180%

results, we can find that a random field with the slower decay rate ($\beta = 0$) is more difficult to be trained when uniform sampling is used. Figure 2 shows that this issue can be resolved by quasi-Monte Carlo sampling with moderate size of training set.

In the literature, asymptotics-based method has been proposed[51] which only works well

for random interfaces with small magnitudes. The proposed method works for random interfaces with large magnitudes. For example, consider $\theta(\omega)$ with 2 random variables ranging over $[-5, 5]$ and $\beta = 2$, the relative L^∞ error is 1.071%; see Figure 10.

For a random field with 10 random variables and 5 realizations $d = [10, 15, 20, 30, 40, 50]$, generalization errors of the trained neural network are plotted in Figures 11, 12, 13 for $\beta = -2, -1, 0$, respectively. A detailed dependence of EDL σ on random variables is given in Figures 14, 15, 16 for $\beta = -2, -1, 0$, respectively.

Contour plot of EDL on random variables in 3D when $\beta = -1$ is given in Figure 17 for comparison.

Supporting Figures

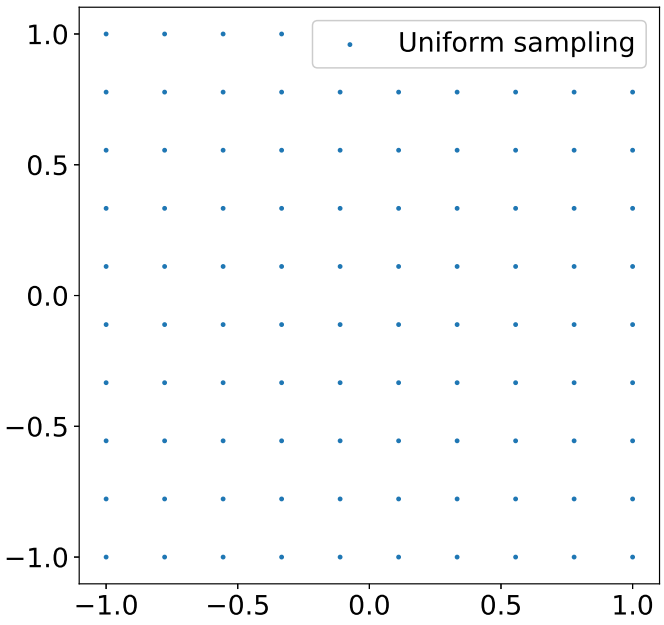


Figure 6: Uniform sampling for two random variables with 100 points.

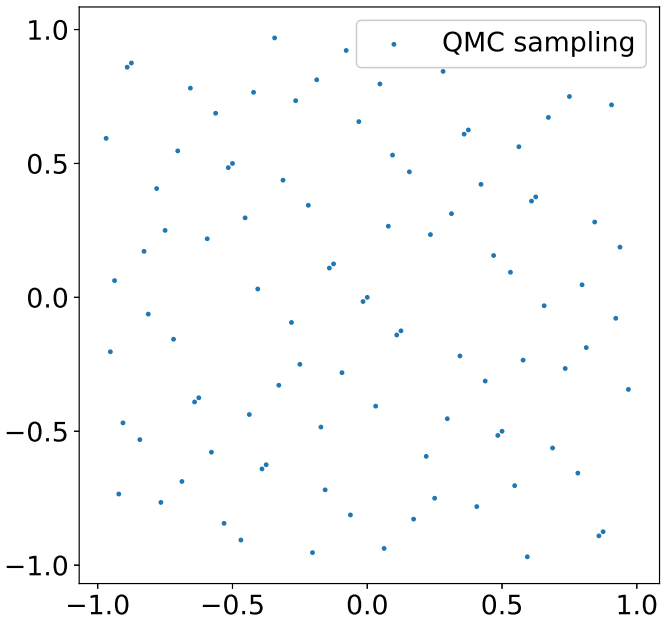


Figure 7: Quasi-Monte Carlo sampling (Sobol sequence) for two random variables with 100 points.

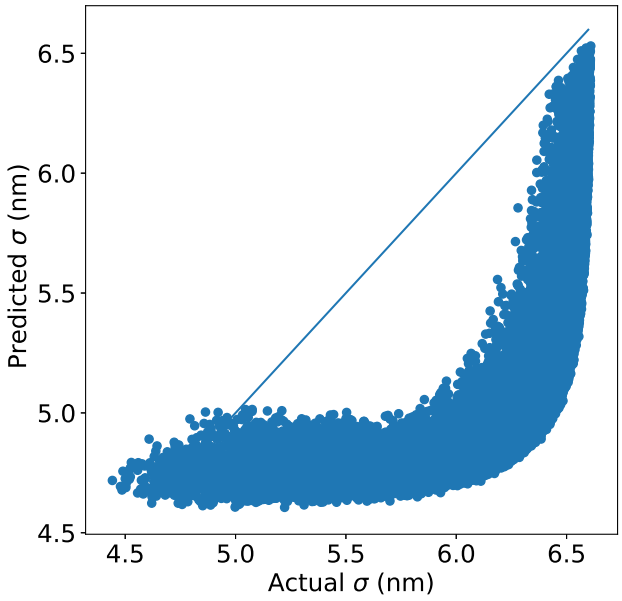


Figure 8: Generalization error of the trained neural network in 2D when the number of random variables is 10 and 1024 uniformly distributed points are used. The relative L^∞ error is 30.561%.

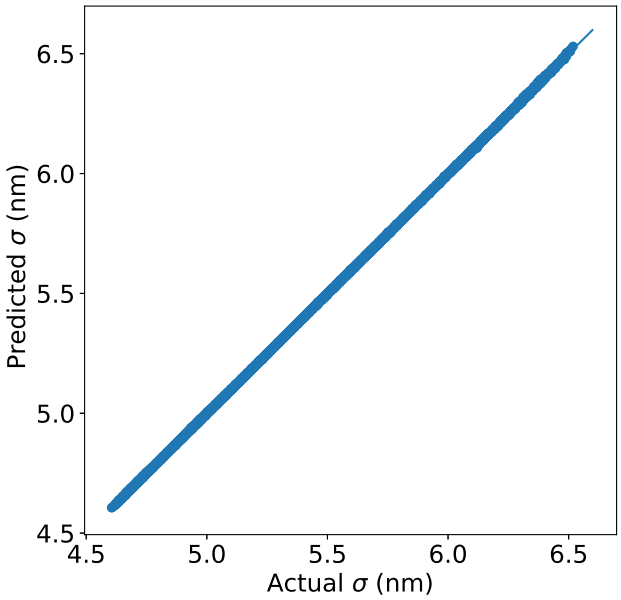


Figure 9: Generalization error of the trained neural network in 2D when the number of random variables is 10 and 1024 points generated by Sobol sequence are used. The relative L^∞ error is 0.237%.

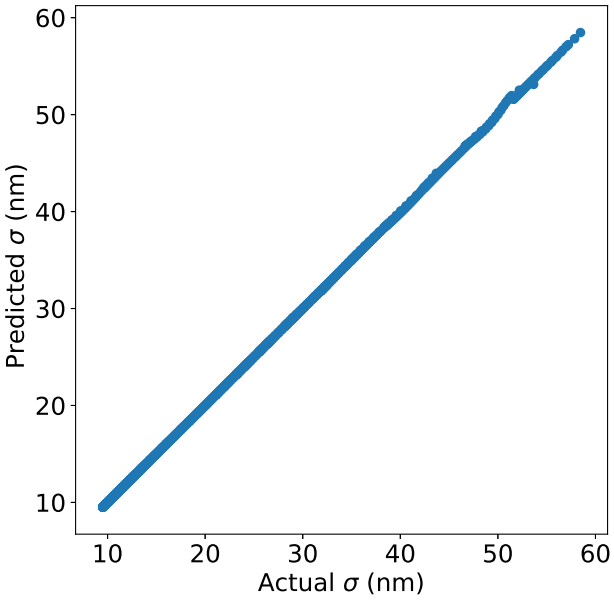


Figure 10: Generalization error of the trained neural network for random variables ranging over $[-5, 5]$ with 1 photoluminescence datum in 2D. The relative L^∞ error is 1.071%.

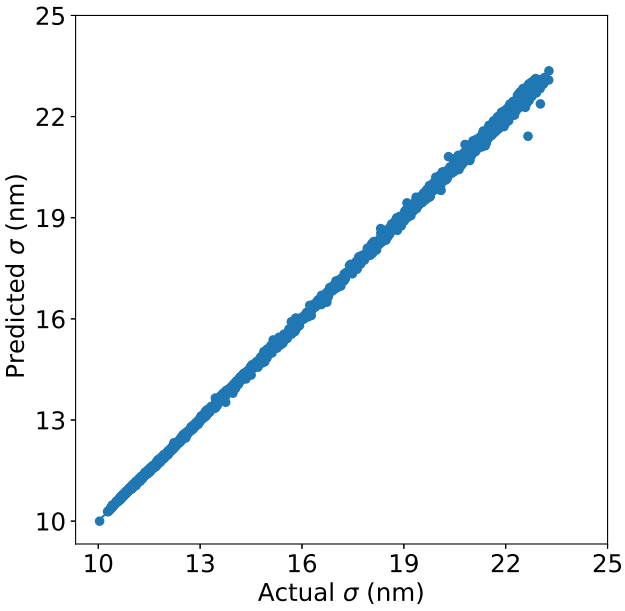


Figure 11: Generalization error of the trained neural network for random variables ranging over $[-5, 5]$ with 6 photoluminescence data and $\beta = -2$ in 2D. The relative L^∞ error is 5.784%.

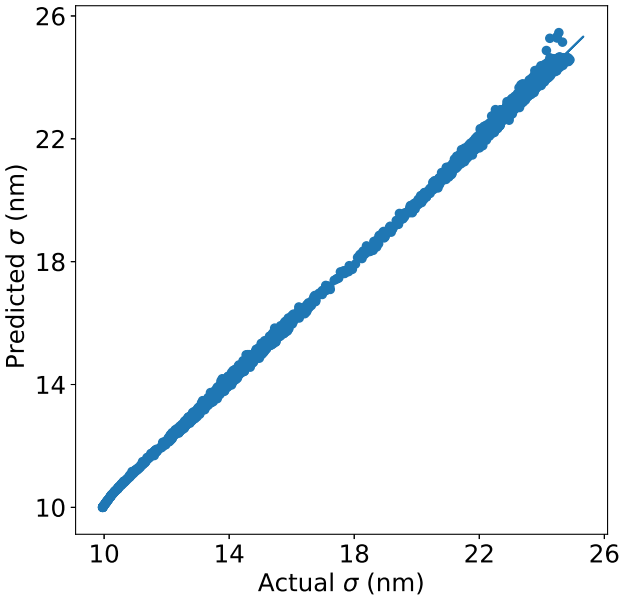


Figure 12: Generalization error of the trained neural network for random variables ranging over $[-5, 5]$ with 6 photoluminescence data and $\beta = -1$ in 2D. The relative L^∞ error is 3.327%.

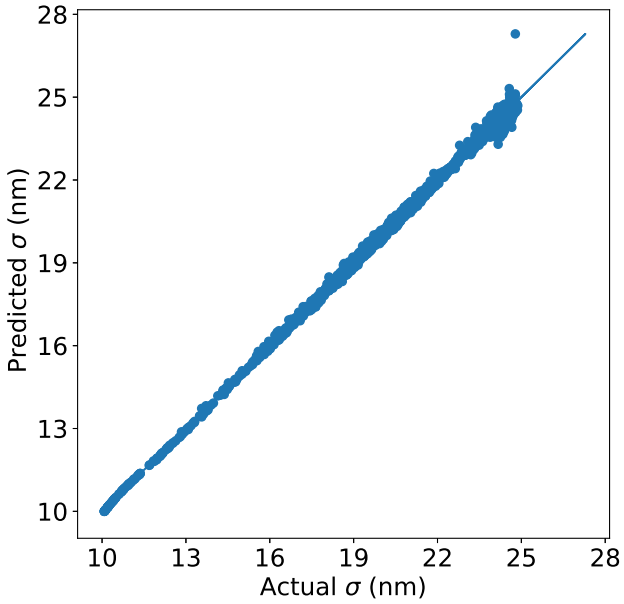


Figure 13: Generalization error of the trained neural network for random variables ranging over $[-5, 5]$ with 6 photoluminescence data and $\beta = 0$ in 2D. The relative L^∞ error is 9.184%.

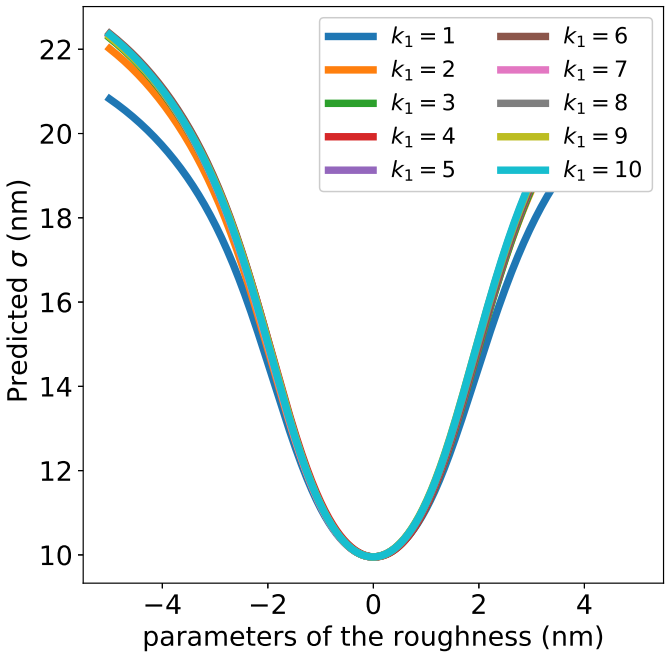


Figure 14: Dependence of exciton diffusion length on random variables in 2D when $\beta = -2$.

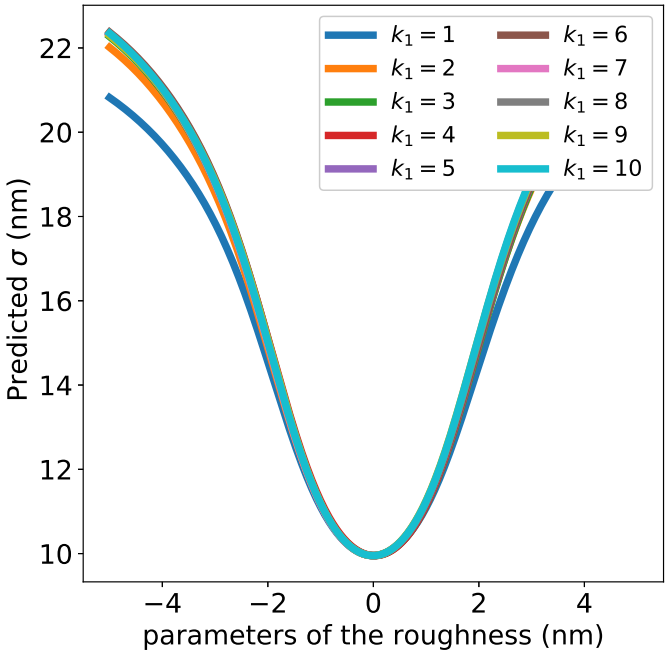


Figure 15: Dependence of exciton diffusion length on random variables in 2D when $\beta = -1$.

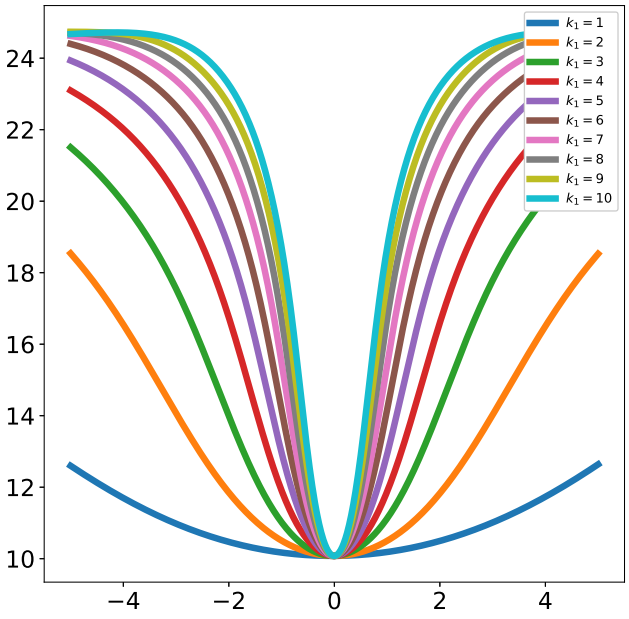


Figure 16: Dependence of exciton diffusion length on random variables in 2D when $\beta = 0$.

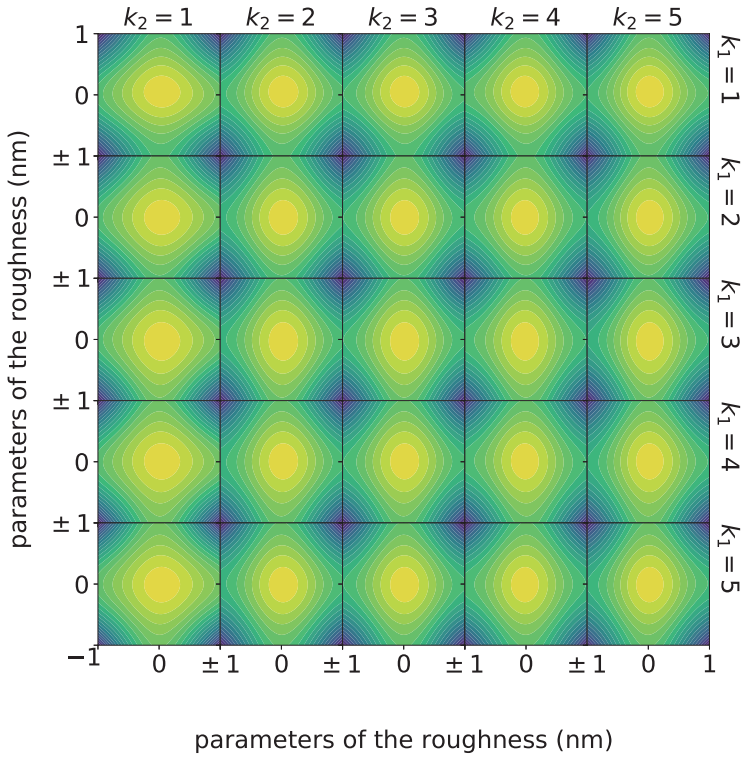


Figure 17: Contour plot of exciton diffusion length on random variables in 3D when $\beta = -1$.

References

(1) Myers, J. D.; Xue, J. Organic Semiconductors and their Applications in Photovoltaic Devices. Polymer Reviews **2012**, 52, 1–37.

(2) Su, Y.-W.; Lan, S.-C.; Wei, K.-H. Organic photovoltaics. Materials Today **2012**, 15, 554 – 562.

(3) Martin, P.; E, S. C. Electronic processes in organic crystals and polymers, 2nd ed.; New York (N.Y.) : Oxford university press, 1999.

(4) Forrest, S. The path to ubiquitous and low-cost organic electronic appliances on plastic. Nature **2004**, 428, 911–918.

(5) Peumans, P.; Yakimov, A.; Forrest, S. Small molecular weight organic thin-film photodetectors and solar cells. Journal of Applied Physics **2003**, 93, 3693–3723.

(6) Lin, J.; Mikhnenko, O.; Chen, J.; Masri, Z.; Ruseckas, A.; Mikhailovsky, A.; Raab, R.; Liu, J.; Blom, P.; Loi, M.; García-Cervera, C.; Samuel, I.; Nguyen, T.-Q. Systematic study of exciton diffusion length in organic semiconductors by six experimental methods. Materials Horizons **2014**, 1, 280–285.

(7) Sanaur, S.; Whalley, A.; Alameddine, B.; Carnes, M.; Nuckolls, C. Jet-printed electrodes and semiconducting oligomers for elaboration of organic thin-film transistors. Organic Electronics **2006**, 7, 423–427.

(8) Murphy, A. R.; Fréchet, J. M. J. Organic Semiconducting Oligomers for Use in Thin Film Transistors. Chemical Reviews **2007**, 107, 1066–1096.

(9) Pettersson, L.; Roman, L.; Inganäs, O. Modeling photocurrent action spectra of photovoltaic devices based on organic thin films. Journal of Applied Physics **1999**, 86, 487–496.

(10) Mikhnenko, O.; Azimi, H.; Scharber, M.; Morana, M.; Blom, P.; Loi, M. Exciton diffusion length in narrow bandgap polymers. Energy and Environmental Science **2012**, 5, 6960–6965.

(11) De Oliveira Neto, P.; Teixeira, J.; Da Cunha, W.; Gargano, R.; E Silva, G. Electron-lattice coupling in armchair graphene nanoribbons. Journal of Physical Chemistry Letters **2012**, 3, 3039–3042.

(12) Brabec, C.; Hauch, J.; Schilinsky, P.; Waldauf, C. Production aspects of organic photovoltaics and their impact on the commercialization of devices. MRS Bulletin **2005**, 30, 50–52.

(13) Brédas, J.-L.; Norton, J.; Cornil, J.; Coropceanu, V. Molecular understanding of organic solar cells: The challenges. Accounts of Chemical Research **2009**, 42, 1691–1699.

(14) Terao, Y.; Sasabe, H.; Adachi, C. Correlation of hole mobility, exciton diffusion length, and solar cell characteristics in phthalocyanine/fullerene organic solar cells. Applied Physics Letters **2007**, 90, 103515.

(15) Menke, S.; Luhman, W.; Holmes, R. Tailored exciton diffusion in organic photovoltaic cells for enhanced power conversion efficiency. Nature Materials **2013**, 12, 152–157.

(16) Antoniadis,; Rothberg,; Papadimitrakopoulos,; Yan, W.; Galvín,; Abkowitz, Enhanced carrier photogeneration by defects in conjugated polymers and its mechanism. Physical review. B, Condensed matter **1994**, 50 20, 14911–14915.

(17) Mullenbach, T.; Curtin, I.; Zhang, T.; Holmes, R. Probing dark exciton diffusion using photovoltage. Nature Communications **2017**, 8, 14215.

(18) Tang, C. Two-layer organic photovoltaic cell. Applied Physics Letters **1986**, 48, 183–185.

(19) Rim, S.-B.; Fink, R.; Schöneboom, J.; Erk, P.; Peumans, P. Effect of molecular packing on the exciton diffusion length in organic solar cells. Applied Physics Letters **2007**, 91, 173504.

(20) Markov, D.; Tanase, C.; Blom, P.; Wildeman, J. Simultaneous enhancement of charge transport and exciton diffusion in poly(p-phenylene vinylene) derivatives. Physical Review B - Condensed Matter and Materials Physics **2005**, 72, 045217.

(21) Scully, S.; McGehee, M. Effects of optical interference and energy transfer on exciton diffusion length measurements in organic semiconductors. Journal of Applied Physics **2006**, 100, 034907.

(22) Goh, C.; Scully, S.; McGehee, M. Effects of molecular interface modification in hybrid organic-inorganic photovoltaic cells. Journal of Applied Physics **2007**, 101, 114503.

(23) Shaw, P.; Ruseckas, A.; Samuel, I. Exciton diffusion measurements in poly(3-hexylthiophene). Advanced Materials **2008**, 20, 3516–3520.

(24) Wu, Y.; Zhou, Y.; Wu, H.; Zhan, Y.; Zhou, J.; Zhang, S.; Zhao, J.; Wang, Z.; Ding, X.; Hou, X. Metal-induced photoluminescence quenching of tri-(8-hydroxyquinoline) aluminum. Applied Physics Letters **2005**, 87.

(25) Theander, M.; Yartsev, A.; Zigmantas, D.; Sundström, V. Photoluminescence quenching at a heterojunction. Physical Review B - Condensed Matter and Materials Physics **2000**, 61, 12957–12963.

(26) Mikhnenko, O.; Lin, J.; Shu, Y.; Anthony, J.; Blom, P.; Nguyen, T.-Q.; Loi, M. Effect of thermal annealing on exciton diffusion in a diketopyrrolopyrrole derivative. Physical Chemistry Chemical Physics **2012**, 14, 14196–14201.

(27) Cook, S.; Furube, A.; Katoh, R.; Han, L. Estimate of singlet diffusion lengths in PCBM films by time-resolved emission studies. Chemical Physics Letters **2009**, 478, 33–36.

(28) Lewis, A.; Ruseckas, A.; Gaudin, O.; Webster, G.; Burn, P.; Samuel, I. Singlet exciton diffusion in MEH-PPV films studied by exciton-exciton annihilation. Organic Electronics **2006**, 7, 452–456.

(29) Masri, Z.; Ruseckas, A.; Emelianova, E.; Wang, L.; Bansal, A.; Matheson, A.; Lemke, H.; Nielsen, M.; Nguyen, H.; Coulembier, O.; Dubois, P.; Beljonne, D.; Samuel, I. Molecular weight dependence of exciton diffusion in poly(3-hexylthiophene). Advanced Energy Materials **2013**, 3, 1445–1453.

(30) Shaw, P.; Ruseckas, A.; Peet, J.; Bazan, G.; Samuel, I. Exciton - Exciton annihilation in mixed-phase polyfluorene films. Advanced Functional Materials **2010**, 20, 155–161.

(31) Qin, D.; Gu, P.; Dhar, R.; Razavipour, S.; Ban, D. Measuring the exciton diffusion length of C 60 in organic planar heterojunction solar cells. Physica Status Solidi (A) Applications and Materials Science **2011**, 208, 1967–1971.

(32) Stübinger, T.; Brütting, W. Exciton diffusion and optical interference in organic donor-acceptor photovoltaic cells. Journal of Applied Physics **2001**, 90, 3632–3641.

(33) Yang, C.; Tang, Z.; Ge, W.; Wang, J.; Zhang, Z.; Jian, X. Exciton diffusion in light-emitting organic thin films studied by photocurrent spectra. Applied Physics Letters **2003**, 83, 1737–1739.

(34) Halls, J.; Pichler, K.; Friend, R.; Moratti, S.; Holmes, A. Exciton diffusion and dissociation in a poly(p-phenylenevinylene)/C60 heterojunction photovoltaic cell. Applied Physics Letters **1996**, 68, 3120–3122.

(35) Ghosh, A.; Feng, T. Merocynanine organic solar cells. Journal of Applied Physics **1978**, 49, 5982–5989.

(36) Wagner, J.; Fritz, T.; Böttcher, H. Computer Modelling of Organic Thin Film Solar

Cells. I. Exciton Model of Photocurrent Generation. Physica Status Solidi (a) **1993**, 136, 423–432.

(37) Bulović, V.; Forrest, S. Excitons in crystalline thin films of 3,4,9,10-perylenetetracarboxylic dianhydride studied by photocurrent response. Chemical Physics Letters **1995**, 238, 88–92.

(38) Rim, S.-B.; Peumans, P. The effects of optical interference on exciton diffusion length measurements using photocurrent spectroscopy. Journal of Applied Physics **2008**, 103, 124515.

(39) Huijser, A.; Savenije, T.; Shalav, A.; Siebbeles, L. An experimental study on the molecular organization and exciton diffusion in a bilayer of a porphyrin and poly(3-hexylthiophene). Journal of Applied Physics **2008**, 104, 034505.

(40) Fravventura, M.; Hwang, J.; Suijkerbuijk, J.; Erk, P.; Siebbeles, L.; Savenije, T. Determination of singlet exciton diffusion length in thin evaporated C 60 films for photovoltaics. Journal of Physical Chemistry Letters **2012**, 3, 2367–2373.

(41) Savenije, T.; Siebbeles, L. Laser Science, Optical Society of America **2010**,

(42) Kroeze, J.; Savenije, T.; Vermeulen, M.; Warman, J. Contactless determination of the photoconductivity action spectrum, exciton diffusion length, and charge separation efficiency in polythiophene-sensitized TiO2 bilayers. Journal of Physical Chemistry B **2003**, 107, 7696–7705.

(43) Lunt, R.; Benziger, J.; Forrest, S. Relationship between crystalline order and exciton diffusion length in molecular organic semiconductors. Advanced Materials **2010**, 22, 1233–1236.

(44) Bergemann, K.; Forrest, S. Measurement of exciton diffusion lengths in optically thin organic films. Applied Physics Letters **2011**, 99, 243303.

(45) Rand, B.; Cheyns, D.; Vasseur, K.; Giebink, N.; Mothy, S.; Yi, Y.; Coropceanu, V.; Beljonne, D.; Cornil, J.; Brédas, J.-L.; Genoe, J. The impact of molecular orientation on the photovoltaic properties of a phthalocyanine/fullerene heterojunction. Advanced Functional Materials **2012**, 22, 2987–2995.

(46) Lunt, R.; Giebink, N.; Belak, A.; Benziger, J.; Forrest, S. Exciton diffusion lengths of organic semiconductor thin films measured by spectrally resolved photoluminescence quenching. Journal of Applied Physics **2009**, 105, 053711.

(47) Mullenbach, T.; McGarry, K.; Luhman, W.; Douglas, C.; Holmes, R. Connecting molecular structure and exciton diffusion length in rubrene derivatives. Advanced Materials **2013**, 25, 3689–3693.

(48) Chen, J.; Lin, J.; Nguyen, T.-Q. Towards a Unified Macroscopic Description of Exciton Diffusion in Organic Semiconductors. Communications in Computational Physics **2016**, 20, 754–772.

(49) Hong, L.; Hu, Z.; Liu, G. Monte carlo methods for value-at-risk and conditional value-at-risk: A review. ACM Transactions on Modeling and Computer Simulation **2014**, 24, 5.

(50) Xiu, D.; Hesthaven, J. S. High-Order Collocation Methods for Differential Equations with Random Inputs. SIAM J. Scientific Computing **2005**, 27, 1118–1139.

(51) Chen, J.; Lin, L.; Zhang, Z.; Zhou, X. Estimation of exciton diffusion lengths of organic semiconductors in random domains. Journal of Computational Physics **2019**, 376, 894–912.

(52) Oliynyk, A.; Adutwum, L.; Rudyk, B.; Pisavadia, H.; Lotfi, S.; Hlukhyy, V.; Harynuk, J.; Mar, A.; Brgoch, J. Disentangling Structural Confusion through Machine Learning: Structure Prediction and Polymorphism of Equiatomic Ternary Phases ABC. Journal of the American Chemical Society **2017**, 139, 17870–17881.

(53) Oliynyk, A.; Adutwum, L.; Harynuk, J.; Mar, A. Classifying crystal structures of binary compounds AB through cluster resolution feature selection and support vector machine analysis. Chemistry of Materials **2016**, 28, 6672–6681.

(54) Hautier, G.; Fischer, C.; Jain, A.; Mueller, T.; Ceder, G. Finding natures missing ternary oxide compounds using machine learning and density functional theory. Chemistry of Materials **2010**, 22, 3762–3767.

(55) Isayev, O.; Oses, C.; Toher, C.; Gossett, E.; Curtarolo, S.; Tropsha, A. Universal fragment descriptors for predicting properties of inorganic crystals. Nature Communications **2017**, 8, 15679.

(56) Xie, T.; Grossman, J. Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties. Physical Review Letters **2018**, 120, 145301.

(57) De Jong, M.; Chen, W.; Notestine, R.; Persson, K.; Ceder, G.; Jain, A.; Asta, M.; Gamst, A. A Statistical Learning Framework for Materials Science: Application to Elastic Moduli of k-nary Inorganic Polycrystalline Compounds. Scientific Reports **2016**, 6, 34256.

(58) Wang, T.; Wu, D. J.; Coates, A.; Ng, A. Y. End-to-end text recognition with convolutional neural networks. Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012) **2012**, 3304–3308.

(59) Krizhevsky, A.; Ilya, S.; E, H. G. In Advances in Neural Information Processing Systems 25; Pereira, F., Burges, C. J. C., Bottou, L., Weinberger, K. Q., Eds.; Curran Associates, Inc., 2012; pp 1097–1105.

(60) Ruhi, S.; E., H. G.; Anoop, D. Application of Deep Belief Networks for Natural Language Understanding. IEEE/ACM Trans. Audio, Speech and Lang. Proc. **2014**, 22, 778–784.

(61) Caflisch, R. Monte Carlo and quasi-Monte Carlo methods; Cambridge University Press, 1998; Vol. 7; pp 1–49.

(62) He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. CoRR **2015**, 1512.03385.

(63) Burkhard, G. F.; Hoke, E. T.; McGehee, M. D. Accounting for Interference, Scattering, and Electrode Absorption to Make Accurate Internal Quantum Efficiency Measurements in Organic and Other Thin Solar Cells. Advanced Materials **2010**, 22, 3293–3297.

(64) Guide, M.; Lin, J. D. A.; Proctor, C. M.; Chen, J.; Garcia-Cervera, C.; Nguyen, T.-Q. Effect of copper metalation of tetrabenzoporphyrin donor material on organic solar cell performance. J. Mater. Chem. A **2014**, 2, 7890–7896.

(65) Nocedal, J.; Wright, S. J. Numerical Optimization; Springer Series in Operations Research, Springer-Verlag: New York, 1999.

(66) Bezanson, J.; Edelman, A.; Karpinski, S.; Shah, V. Julia: A fresh approach to numerical computing. SIAM Review **2017**, 59, 65–98.

Graphical TOC Entry