



OPEN

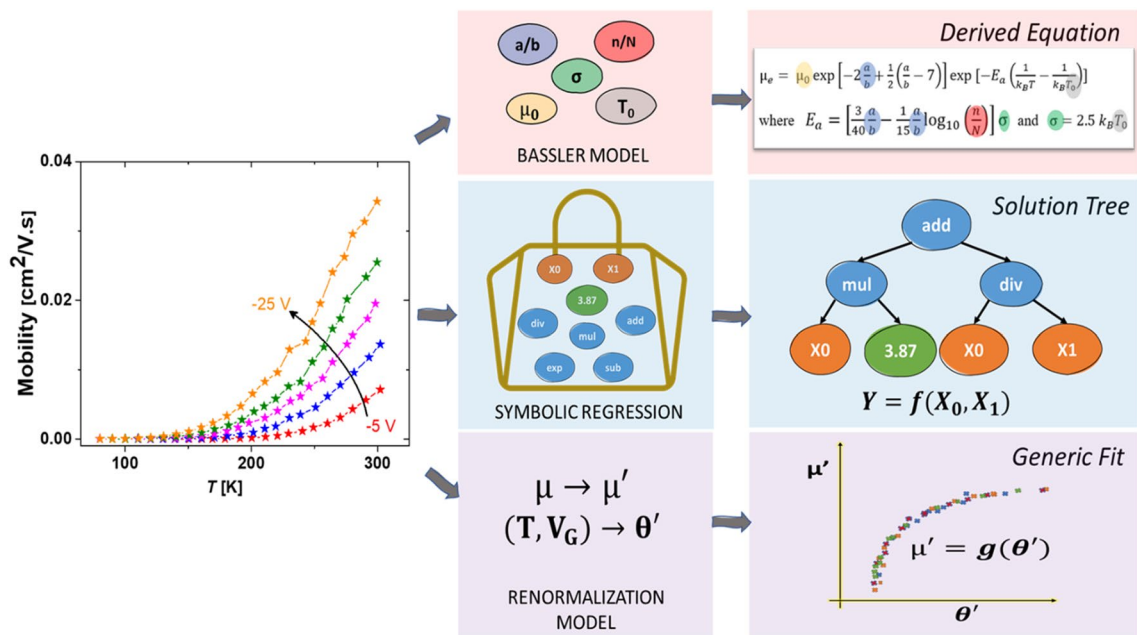
## Comparing data driven and physics inspired models for hopping transport in organic field effect transistors

Madhavkrishnan Lakshminarayanan<sup>1,2,5</sup>, Rajdeep Dutta<sup>3,5</sup>, D. V. Maheswar Repaka<sup>2</sup>, Senthilnath Jayavelu<sup>3</sup>✉, Wei Lin Leong<sup>1</sup>✉ & Kedar Hippalgaonkar<sup>2,4</sup>✉

The past few decades have seen an uptick in the scope and range of device applications of organic semiconductors, such as organic field-effect transistors, organic photovoltaics and light-emitting diodes. Several researchers have studied electrical transport in these materials and proposed physical models to describe charge transport with different material parameters, with most disordered semiconductors exhibiting hopping transport. However, there exists a lack of a consensus among the different models to describe hopping transport accurately and uniformly. In this work, we first evaluate the efficacy of using a purely data-driven approach, i.e., symbolic regression, in unravelling the relationship between the measured field-effect mobility and the controllable inputs of temperature and gate voltage. While the regressor is able to capture the scaled mobility well with mean absolute error (MAE)  $\sim O(10^{-2})$ , better than the traditionally used hopping transport model, it is unable to derive physically interpretable input–output relationships. We then examine a physics-inspired renormalization approach to describe the scaled mobility with respect to a scale-invariant reference temperature. We observe that the renormalization approach offers more generality and interpretability with a MAE of the  $\sim O(10^{-1})$ , still better than the traditionally used hopping model, but less accurate as compared to the symbolic regression approach. Our work shows that physics-based approaches are powerful compared to purely data-driven modelling, providing an intuitive understanding of data with extrapolative ability.

Several classes of organic semiconductors—conjugated polymers and small molecules—have been discovered and explored in a wide range of electronic applications, such as flexible displays<sup>1</sup>, self-healing materials for electronic skin<sup>2</sup>, energy harvesting<sup>3,4</sup>, chemical and biological sensors<sup>5</sup>. The electrical transport of these semiconductors have been studied using time-of-flight (TOF) or charge extraction by linearly increasing voltage (CELIV) measurements<sup>6</sup> and most commonly by fabricating organic field-effect transistors (OFETs) employed in Hall measurements<sup>7–9</sup>. The electrical transport in these materials is complex and no specific transport mechanism has been established due to the weak interactions between molecules and strong dependence on processing conditions and device morphologies. Such weak van der Waals interactions lead to different degrees of disorder ranging from amorphous to crystalline materials, which affects the electrical transport. Small molecules and single crystals in which  $\pi$ – $\pi$  interactions dominate possess less degree of disorder and are expected to follow band-like transport<sup>10</sup>. Alternatively, polymers that lack order in packing are thereby expected to follow thermally-driven hopping transport<sup>11</sup>. Thus, thermally activated Arrhenius-type hopping usually describes the electrical transport in conjugated polymers with the assumption that a fraction of the total number of available energy states is occupied. While physical models like the Miller–Abrahams formalism<sup>12</sup>, explaining hopping transport in material systems, are complex involving many parameters that are defined empirically, it provides an intuitive

<sup>1</sup>School of Electrical Electronic Engineering, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798, Singapore. <sup>2</sup>Institute of Materials Research & Engineering, Agency for Science, Technology and Research (A\*STAR), Singapore 138632, Singapore. <sup>3</sup>Institute for Infocomm Research, Agency for Science, Technology and Research (A\*STAR), Singapore 138632, Singapore. <sup>4</sup>School of Materials Science and Engineering, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798, Singapore. <sup>5</sup>These authors contributed equally: Madhavkrishnan Lakshminarayanan and Rajdeep Dutta. ✉email: j\_senthilnath@i2r.a-star.edu.sg; wleong@ntu.edu.sg; kedar@ntu.edu.sg

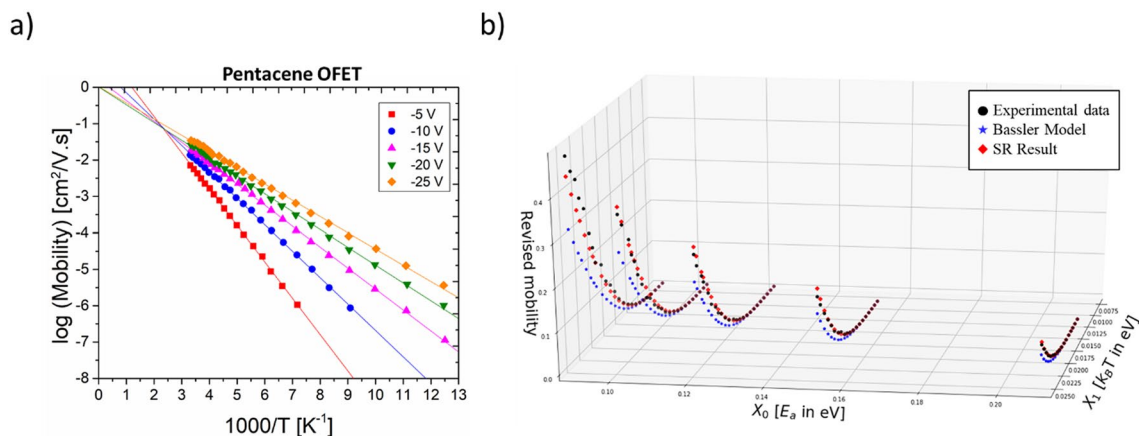


**Figure 1.** An overview of various approaches to extract the underlying input–output relationships from temperature-dependent field-effect mobility data. Here, the left panel consists of a representative temperature-dependent field-effect mobility data at different gate voltages. The middle panel captures the methods followed in each approach. The panel on the right shows the final solution form for each approach. For the Bassler Model, the unknowns are coloured, and the derived equation fits these values. In Symbolic Regression, inputs ( $X_0$  and  $X_1$ ), constants and mathematical operations are fed into the algorithm, while a solution tree produces a fitting function,  $Y$ . For the Renormalization Model, a generic fit by rescaling the inputs is obtained.

understanding of the physics of transport and hence can be predictive in nature. In essence, this boils down to the dependence of charge hopping upon the thermodynamic temperature ( $T$ ) and the modulation of additional carriers in the polymeric channel via the applied gate voltage ( $V_g$ ) in the OFET configuration.

Similar efforts at understanding physical principles in other materials and physical systems have been employed in recent times with simulated physics datasets. “AI Feynman” (AI stands for Artificial Intelligence) is a recursive symbolic regression algorithm, which connects neural networks with physics-inspired methods such as dimensionality analysis, generalized symmetries, multiplicative separability, and inversion to discover physical equations that explain the data<sup>13,14</sup>. The use of symbolic neural networks to learn the dynamics of data with a Partial Differential Equation (PDE-NET) fixes the functional form but is able to capture the underlying parameters<sup>15</sup>. Physics-informed Neural Networks (PINNs) respect the laws of physics while solving both differentiable (data-efficient approximators) and discrete datasets and have been used to solve problems in fluids<sup>16</sup>, quantum mechanics and more<sup>17</sup>. In this work, we investigate the efficacy of a powerful mathematical approach, namely symbolic regression (SR), in recapturing underlying physics after fitting transport data of three different OFETs based on organic molecules of p-type pentacene, and n-type buckminsterfullerene ( $C_{60}$ ) as well as conjugated polymer poly(3-hexylthiophene-2,5-diyl) (P3HT). While the regressor can fit the experimental data after carefully tuning the hyperparameters, physical interpretability is not easy and guaranteed in all cases. The regressor does succeed in identifying the physically meaningful and essential terms to describe the transport. However, it is unable to decouple the effect of the activation energy modulated by varying gate voltage from the magnitude of the thermal energy window that is directly proportional to temperature. We see that the regressor treats the whole dataset as one entity, as opposed to a system of five curves following the same bi-parametric equation, where the data in each curve is generated by only varying the value of one of the parameters, while keeping the other parameter constant. The resulting equation fits the entire dataset at once instead of providing a solution that is repeatable curve-by-curve by varying the value of the second variable/parameter (gate voltage in this case). To overcome this missing physical insight, we compare this with a renormalization approach to describe the transport by defining dimensionless quantities from the temperature dependent mobility data. Derivation of dimensionless quantities captures both the effects of temperature as well as gate voltage embedded in a single parameter, rather than in multiple parameters as seen in the existing models<sup>18</sup>, and provides us with a one-shot approach in describing and characterising hopping transport in organic semiconductors. We demonstrate that this approach is powerful as it also helps to differentiate between processing conditions for the same molecules by comparing the extent of activation energy that is accessed over the range of the applied gate voltages. Figure 1 depicts various approaches for extracting input–output relationships involved in hopping transport phenomena.

As a benchmark, we use the model derived by Bassler et al. for the field-effect mobility ( $\mu_e$ ) using the effective medium approximation and the concept of effective transport energy<sup>11,19</sup>, while considering Gaussian density of states (DOS) and Miller–Abrahams jump rates for thermally activated Arrhenius-type hopping. Here,



**Figure 2.** (a) Temperature-dependent mobility data for pentacene OFET; (b) Simultaneous visualization of the fits obtained from symbolic regression (SR) (red symbol) and the Bassler model (blue symbol) with respect to the experimental data (black symbol) (data extracted from Ref.<sup>19</sup>).

according to the so-called Bassler model,  $\mu_e = \mu_0 \exp\left[-2\frac{a}{b} + \frac{1}{2}\left(\frac{a}{b} - 7\right)\right] \exp\left[-E_a\left(\frac{1}{k_B T} - \frac{1}{k_B T_0}\right)\right]$ , where  $E_a = \left[\frac{3}{40}\frac{a}{b} - \frac{1}{15}\frac{a}{b}\log_{10}\left(\frac{n}{N}\right)\right]\sigma$  and  $\sigma = 2.5k_B T_0$ .

As illustrated in Fig. S3,  $\sigma$  indicates the broadening of the Gaussian DOS,  $a/b$  is the ratio of inter-site distance and localization radius. Further,  $n/N$  is the ratio of occupied and total number of transport states.  $T_0$  is the isokinetic temperature at which all curves of varying  $E_a$  converge;  $n$  is the number of occupied transport states, which can be tuned by the applied gate voltage ( $V_g$ ) since it is directly proportional to  $V_g$ . The total number of transport states,  $N$ , is a constant of the order of  $\sim 10^{23} \text{ cm}^{-3}$ , which relates to the lattice constant ( $a$ ) as  $N = a^{-3}$  for an assumed cubic lattice.

In the equation above, the hopping process is affected by both spatial ( $a$ ,  $b$ ) and energetic terms ( $E_a$ ). The spatial terms indicate the extent of hopping (given by  $a$ ) and the localization radius between two sites (given by  $b$ ), while the energetic terms specify the activation required for the phonon-assisted excitation, spatial tunnelling, and phonon emission—all together considered as one hopping event—to take place from one site to another. The tunnelling strongly depends on the extent of wave function overlap between neighbouring molecular orbitals. This results in an increase in mobility when the wave functions overlap is higher, which thereby decreases the activation energy for the hopping event<sup>20,21</sup>. The activation energy also relates to the structure and morphology of the semiconducting layer through molecular orientation and processing conditions.

### Data description

The temperature-dependent mobility data for pentacene, P3HT,  $C_{60}$  OFETs is obtained from literature<sup>22–24</sup>. While pentacene and P3HT were fabricated through solution processing methods on Si/SiO<sub>2</sub> substrates with Au contacts,  $C_{60}$  film was fabricated through standard flush evaporation or hot-wall epitaxy<sup>25</sup> on ITO/divinyltetramethyldisiloxane-bis(benzocyclobutane) (BCB) substrates with LiF/Al top contacts. The pentacene film can be regarded as polycrystalline, while the P3HT film has ordered regions that alternate with disordered regions. The  $C_{60}$  film is of greater crystallinity compared to both P3HT and pentacene films. The field-effect mobility (output) data recorded as a function of the change in two variables namely temperature and gate voltage (inputs), is shown in Fig. 2a for pentacene OFET and Fig. S1 for P3HT and  $C_{60}$  OFETs. Typical of materials science dataset sizes, each OFET data is comprised of  $\sim 100$  points, stacking input–output information from five temperature-dependent curves at five different gate voltages that were used to record the data (as seen in Fig. 2a). The activation energy can be regulated by varying the applied gate voltage, i.e., a larger gate voltage signifies that a smaller activation energy is needed.

### Symbolic regression: an approach for data-driven learning of OFET behaviour

First, we introduce Symbolic Regression (SR) and discuss the fundamentals of genetic programming and how it is employed specifically in this work to extract the underlying input–output relationship in the form of comprehensive mathematical expressions<sup>26,27</sup>. Although other machine learning methods are useful in extracting input–output relations, they employ predetermined functional forms to fit the data (for instance, polynomial regression). On the contrary, SR attempts to solve such problems without requiring any prior knowledge of the functional forms, by utilizing a finite set of operators and functions (like arithmetic, trigonometric, exponential). Such a symbolic input–output mapping is beneficial as it offers interpretability in applied physics and material science, which is the motivation behind using this technique over other machine learning models<sup>27–29</sup>.

For characterizing hopping-like transport phenomena in OFETs, several models have been proposed to relate the field-effect mobility with inputs like temperature and back-gate voltage. However, these models differ in both the functional forms as well as the nature and number of the transport parameters used. In the present work, SR is applied on each OFET dataset where the input variables are defined as  $X_0 = E_a$  and  $X_1 = k_B T$  and the output

Molecule	$\log \mu_e$	MAE from symbolic regression	MAE from Bassler Model	MAE from re-normalization fitting
Pentacene	$-0.52 + 9.37(E_a * k_B T) - 0.38 \frac{E_a}{k_B T} + 8.99E_a - 842.18(k_B T)^2 + 46.64k_B T$	0.027	0.204	0.224
P3HT	$100E_a k_B T - 0.273 \frac{E_a}{k_B T} + 3E_a + 31.9k_B T + 0.138$	0.037	0.145	0.907
C <sub>60</sub>	$-15.31(E_a * k_B T) - 0.47 \frac{E_a}{k_B T} + 15.78E_a - 9.5k_B T + 0.29$	0.0123	0.488	0.0717

**Table 1.** A comparison of the MAE values obtained from the three different fitting approaches in consideration—symbolic regression (SR), Bassler model and Renormalization fitting—for pentacene, P3HT and C<sub>60</sub> OFETs. The second column consists of the resulting equations obtained from the SR fits.

variable as  $Y \propto \log\left(\frac{\mu_e}{\text{prefactor}}\right)$ . The activation energy  $E_a$  for each  $V_g$  curve is obtained from the slope of the respective curve by expressing the logarithmic mobility as a function of the reciprocal temperature. The thermal energy associated at temperature T is calculated as a product of the temperature T and the Boltzmann constant,  $k_B$ . This choice of the two input variables,  $X_0 = E_a$  and  $X_1 = k_B T$  is so that both the variables are in the same units (eV) and hence guide the SR outcome expressions into a meaningful and interpretable form, since the range of values for both  $E_a$  and  $k_B T$  were similar. This is of special significance because a choice of  $E_a \gg k_B T$  would mean that the hopping energy is too large, and transport does not take place. On the other hand,  $E_a \ll k_B T$  would result in a transport mechanism that is different from hopping, akin to metallic conduction, which is unlikely in organic thin films. Further, the usage of dimensionless input variables obtained by normalising the existing  $X_0$  and  $X_1$  parameters with the thermal energy at room temperature, i.e.,  $k_B(300K)$ , was also explored and discarded as it did not help in accuracy or interpretability.

The output variable Y, denoting revised mobility, is a dimensionless quantity. The Bassler model expression can be written in  $\log_{10}$  scale as,

$$\log\left(\frac{\mu_e}{\text{prefactor}}\right) = -\frac{E_a}{2.303} \left(\frac{1}{k_B T} - \frac{1}{k_B T_0}\right). \quad (1)$$

where,  $\frac{\mu_e}{\text{prefactor}}$  is the dimensionless revised mobility.

The framework applied in this paper seeks to find out a mapping  $Y = f(X_0, X_1)$ . We use the *gplearn* package in Python to perform SR<sup>26</sup>.

GP minimizes the Mean Absolute Error (MAE) between the experimental data and the predictions to produce symbolic expressions. In addition, upon evaluating the best values of SR parameters, specifically, the parsimony coefficient and the tournament size (details can be found in Supplementary Information) are determined using grid search, and the resulting trees render mathematical expressions for the three different OFETs as shown in Table 1. After inspecting the results from SR, we find that the outcome expressions fit the experimental data with an MAE of ~0.02–0.03 with a fit better than that offered by the Bassler model; a comparison of the results between the SR outcome and the Bassler Model, is shown in Fig. 2b.

In Table 1, SR outcome expressions capture the expected terms according to the Bassler model, i.e.,  $E_a$  and  $(E_a/k_B T)$ , and include higher-order non-linear terms. Therefore, the coefficients attached to these terms, expectedly do not match the coefficients in Eq. (1)<sup>27</sup>. In summary, the SR gives us better MAE at a cost of lack of generality from a physical equation, which is what the Bassler Model offers. An innate limitation of the SR approach is limited accuracy with sparse data, which can be mitigated by using more densely populated data<sup>13</sup>.

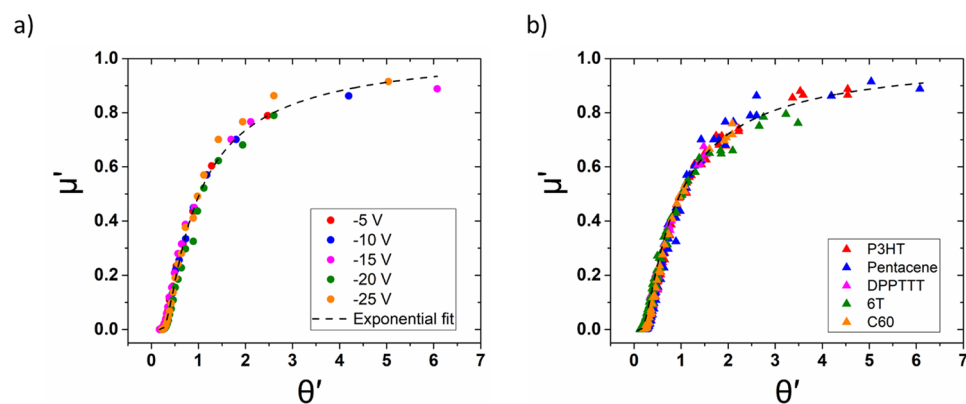
## Physics-inspired model through renormalization

In order to extract underlying physical equations, we next consider a simple yet robust physics-driven approach that makes use of fewer parameters to characterize the underlying transport phenomena in OFETs. In solid-state physics, the concept of renormalization theory is utilized to renormalize the inputs into dimensionless quantities in such a way that certain critical points are identified as scale-invariant, which captures the physics behind the observed data. This has been employed in the study of second-order magnetic phase transitions, for instance<sup>30,31</sup>.

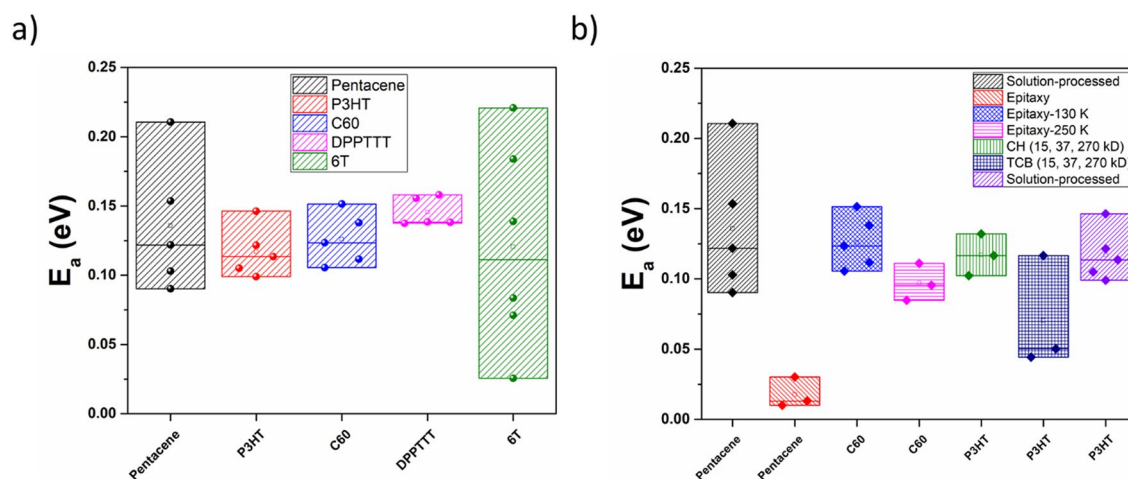
We consider the same OFET data as used for the SR and the Bassler Model. We notice that the carrier mobility, the average velocity of the charge carrier to the electric field, of these disordered systems strongly depends on the carrier concentration, i.e., on applied gate bias ( $\mu \propto V_g^n$ , where  $n$  is the exponent)<sup>32</sup>. The carrier mobility of these organic materials is expected to follow a similar trend with several parameters needed to describe its temperature and field-dependent behaviour. Typically, the carrier mobility of disordered systems can be represented with the help of a scaling law given by  $\mu \propto T^y e^{\frac{E_a}{k_B T}}$ <sup>33</sup>.

In this particular case, for each mobility curve at fixed gate bias ( $V_g$ ),  $T_p$  is the temperature corresponding to the maximum mobility ( $\mu_p$ ) and  $T_r$  is the reference temperature corresponding to half of the maximum mobility ( $\mu_p/2$ ). Due to the limitation of a small number of data points,  $T_r$  is obtained by interpolation directly from the mobility-temperature plots. We choose  $T_p$  as 300 K as the scale-invariant point ( $T_p$ ) in this work for all the studied OFETs.

Upon extracting the  $T_r$  values, we defined two dimensionless quantities  $\mu' = \frac{\mu}{\mu_p}$  and  $\theta' = \frac{T_r - T_p}{T - T_p}$  for every data point in the dataset.  $\mu'$  always takes a value greater than 0 and reaches a maximum value of 1 at  $T = T_p$ , i.e.,  $0 \leq \mu' \leq 1$ . We have obtained the universal plot using temperature-dependent mobility at different  $V_g$ . The



**Figure 3.** (a) Scaled mobility data at different gate voltages (coloured circles) for pentacene OFET and the corresponding exponential curve that fits the scaled data (black dashed line) (data extracted from Ref.<sup>19</sup>); (b) juxtaposition of scaled mobility data at different gate voltages for five different molecules (coloured triangles) resulting in all of them following the same exponential function (shown as black dashed line) as observed in (a).

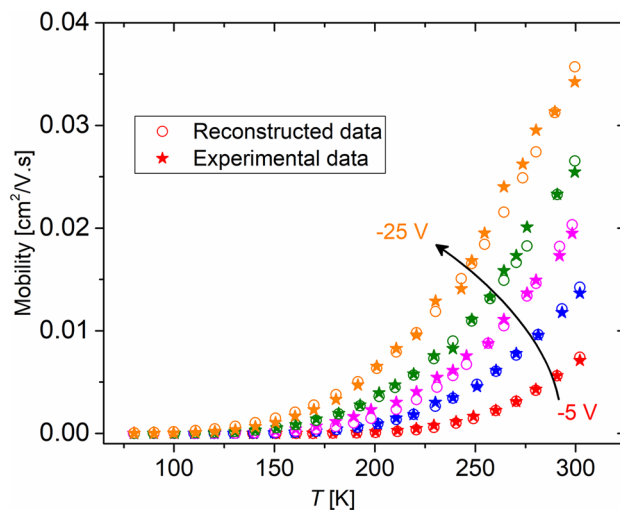


**Figure 4.** (a, b) Plots consisting of activation energy ( $E_a$ ) derived from the Renormalization Model, accessed by different OFETs. The horizontal axes denote the semiconductor for which the activation energy  $E_a$  is plotted on the vertical axes. For each organic system, the plots display the entire range of  $E_a$  values, with the data points being indicated by solid symbols. The minimum and maximum  $E_a$  values are marked by the topmost and bottommost data points, respectively. The corresponding mean and median  $E_a$  values for each OFET are indicated by the unfilled square and horizontal line, respectively, within each box; (a) activation energies ( $E_a$ ) accessed by each OFET as calculated from the range of gate voltages used in calculating the mobility-temperature dependence for each molecule. The  $E_a$  values have been calculated from the values of  $T_r$  and  $T_p$  for each curve and are shown for each OFET (filled spheres). The inset shows the semiconducting material whose  $E_a$  is plotted; (b) variation in the range of  $E_a$  accessed by each OFET fabricated using different processing techniques for the same semiconducting material. The inset represents the different processing conditions used as indicated by varying box patterns.

dimensionless quantities  $\mu' = \frac{\mu}{\mu_p}$  versus  $\theta' = \frac{T_r - T_p}{T - T_p}$  curves collapse into a single curve for different  $V_g$ , as seen for pentacene OFET data in Fig. 3a.

The obtained universal curve is well fitted by an exponential function  $\mu' = a \exp\left(\frac{b}{\theta' + c}\right)$  with  $a \cong 1.04$ ,  $b \cong -0.65$ , and  $c \cong -0.13$  being the fitting constants, respectively. With simple dimensionless normalization parameters we find a comprehensive universality, where all the mobility curves at different gate voltages collapse into a single curve, exhibiting a general form across different organic semiconducting materials of the same hopping mechanism, as evidenced from Fig. 3b.

Furthermore, the knowledge of the  $T_r$  value enables us to calculate the value of the activation energy for that temperature-dependent mobility curve. The activation energy  $E_a$  can be expressed in terms of  $T_p$  and  $T_r$  as  $E_a = k_B \ln(2) \frac{T_r T_p}{T_p - T_r}$ , by evaluating the Bassler model equation at  $T = T_p$  and  $T = T_r$ . Section S1 in the Supplementary Information gives a more detailed derivation of this relationship. Figure 4a shows the obtained/calculated range of activation energies, which are accessible to five different molecules by varying the gate voltage. In general,



**Figure 5.** Comparison between the experimental data (blue symbols) and the reconstructed values from renormalization (red symbols) for pentacene OFET (data extracted from Ref.<sup>19</sup>).

at low gating (low energy state filling) charge carriers require more energy for the hopping to occur, and this activation energy decreases as the gating voltage is increased (more negative values). The strength of this approach, in comparison to SR and the Bassler Model described earlier, is that a general rule to fix  $T_r$  as a function of  $V_g$ , and the values of  $E_a$ , are sufficient to capture the trends for five different organic systems. This, therefore, can be regarded as a generic fitting approach wherein the only descriptor required to convert the  $\mu-T$  data to the  $\mu'-\theta'$  data is the value of  $T_r$  for each  $V_g$  curve.

Figure 4b shows the activation energies accessible to each OFET that were fabricated under different processing conditions. Our results indicate that solution-processed P3HT films with chloroform (CH) require higher activation energies on average than the films grown in trichlorobenzene (TCB)<sup>34</sup>. The mobility is much higher for the TCB films as they require lower activation energy for the hopping transport. This is in line with previous reports on the solution processing of P3HT, wherein the higher mobility of TCB films is attributed to the lower volatility and improved crystallization resulting from deposition by spin-coating<sup>35</sup>.

Similarly, for  $C_{60}$ , the films grown through hot wall epitaxy at  $T_{sub}=250$  °C show a lower band of activation energies accessed for almost similar values of gate voltage as compared to the films grown at  $T_{sub}=130$  °C<sup>24</sup>. This is again consistent with the overall level of crystallization of the resultant films, as evidenced by the AFM images and the value of the broadening of the DOS in either case (75.5 meV at  $T_{sub}=130$  °C and 53.5 meV at  $T_{sub}=250$  °C). A narrow broadening or a sharper DOS is seen in the case of molecules that have better crystallite ordering. Thus, the epitaxially grown pentacene films<sup>36</sup> access a small range of activation energies and are less dispersed, as compared to the solution processed films<sup>22</sup>.

## Discussion

We have shown previously that the activation energy  $E_a$  can be expressed as  $E_a = k_B \ln(2) \frac{T_r T_p}{T_p - T_r}$ . This relation is invertible for all molecules as both the original relation and the universal relation express mobility as an exponential function of the negative of the reciprocal temperature. The original and reconstructed fits for pentacene are shown in Fig. 5.

This approach is not only useful for studying the effect of gate voltage on the activation energy, but also that of the processing conditions such as the choice of solvent, polymer weight, and physical growth temperature, to name a few. This approach could also be extended to other techniques of controlling the extent of energy state filling such as chemical (or vacuum de-doping). Also, we show that the choice of  $T_r$  does not alter the nature of the function that fits the scaled data (only the coefficients  $a$ ,  $b$ ,  $c$  vary). We illustrate this in Section S2 of the SI with the help of temperature-dependent four-probe mobility data of a 2L-pentacene OFET<sup>36</sup>, where mobility at  $T = T_r$  was taken as 70% of  $\mu_p$ . Another salient feature is that the general fit shown in Fig. 3b can also act as a screening technique to determine which OFETs exhibit thermally activated hopping transport, with the help of very few measurements at different temperatures. For example, with mobility measurements at two different temperatures under constant gate voltage, with one at  $T_p$  and another below  $T_p$ , the  $\mu'$  values of these points can be calculated as  $\mu' = \frac{\mu}{\mu_p}$  and subsequently the corresponding  $\theta'$  values can be extracted from the plot (alternatively, the exponential equation) of the universal fit (Fig. 3b). From the  $\theta'$  values thus obtained, the  $T_r$  value of the data point below  $T_p$  can be determined. By virtue of the definition of the scaling relations, upon calculating  $\mu'$  and  $\theta'$  values at  $T = T_p$ , we obtain values of 0.5 and 1 for  $\mu'$  and  $\theta'$ , respectively, if the transport in the OFET occurs through thermally activated hopping. Hence, this not only serves as a screening/sanity check for the nature of charge transport but also helps in evaluating the activation energy at the applied gate voltage with the help of the  $T_r$  value determined thus.

A summary of the different data-fitting approaches is available in Table 2.

SR fitting	Accuracy and interpretability, but no generality. Considers non-linear terms
Bassler model	Physics-based fitting, but material- and range-specific. Not accurate beyond the bounds selected above, so not general
Renormalization fitting	General and interpretable approach but requires rescaling depending on the physics at play. Reconstruction is less accurate compared to Bassler and SR

**Table 2.** A qualitative comparison of the three fitting approaches—Symbolic Regression, Bassler model and Renormalization model, as inferred from our results.

## Conclusions

In this work, we use gate voltage and temperature-dependent OFET data from various small molecules/polymers and show that symbolic regression outperforms physics-inspired fitting techniques such as the Bassler model and the renormalization approach in fitting the OFET data, by one order of magnitude in the corresponding MAE values, although not offering generality achieved through physics-inspired fitting approaches. Through simple graphical value extraction, we demonstrate that the hopping mechanism in OFETs can be described by a general equation that relates normalized mobility with a scaled, dimensionless temperature descriptor. It is challenging to come up with a single model that accurately captures mobility characteristics in both low and high temperature regions; however, the SR outcome and physics-inspired model prove to be better than the baseline in this regard. Although powerful mathematical approaches like symbolic regression may offer reasonably good performance in terms of accuracy of fits, the physical interpretability is often incomplete and not consistent with the established know-how. We also proffer that a similar methodology can be extended to other transport mechanisms wherein the mobility is normalized with respect to the highest observed value and the temperature is then reduced to a dimensionless quantity that encompasses the dependency of the mobility on both the temperature and another experimental variables such as the gate effect or other processing conditions.

Received: 5 August 2021; Accepted: 15 November 2021

Published online: 08 December 2021

## References

- Zhou, L. *et al.* All-organic active matrix flexible display. *Appl. Phys. Lett.* **88**, 083502 (2006).
- Benight, S. J., Wang, C., Tok, J. B. H. & Bao, Z. Stretchable and self-healing polymers and devices for electronic skin. *Prog. Polym. Sci.* **38**, 1961–1977 (2013).
- Arai, R., Furukawa, S., Sato, N. & Yasuda, T. Organic energy-harvesting devices achieving power conversion efficiencies over 20% under ambient indoor lighting. *J. Mater. Chem. A* **7**, 20187–20192 (2019).
- Kaul, A. B. Organic semiconductors and their application in energy harvesting. in *Energy Harvesting and Storage: Materials, Devices, and Applications III*, Vol. 8377 (eds Dhar, N. K. *et al.*) 83770F–83770F–8 (SPIE, 2012).
- Lin, P. & Yan, F. Organic thin-film transistors for chemical and biological sensing. *Adv. Mater.* **24**, 34–51 (2012).
- Stephen, M., Genevicius, K., Juška, G., Arlauskas, K. & Hiorns, R. C. Charge transport and its characterization using photo-CELIV in bulk heterojunction solar cells. *Polym. Int.* **66**, 13–25 (2017).
- Uemura, T. *et al.* Temperature dependence of the Hall effect in pentacene field-effect transistors: Possibility of charge decoherence induced by molecular fluctuations. *Phys. Rev. B Condens. Matter Mater. Phys.* **85**, 035313 (2012).
- Wang, S., Ha, M., Manno, M., Daniel Frisbie, C. & Leighton, C. Hopping transport and the Hall effect near the insulator-metal transition in electrochemically gated poly(3-hexylthiophene) transistors. *Nat. Commun.* **3**, 1–7 (2012).
- Chang, J. F. *et al.* Hall-effect measurements probing the degree of charge-carrier delocalization in solution-processed crystalline molecular semiconductors. *Phys. Rev. Lett.* **107**, 066601 (2011).
- Lu, N. & Geng, D. A review for polaron dependent charge transport in organic semiconductor. *Org. Electron.* **61**, 223–234 (2018).
- Fishchuk, I. I., Arkhipov, V. I., Kadashchuk, A. & Heremans, P. Analytic model of hopping mobility at large charge carrier concentrations in disordered organic semiconductors: Polarons versus bare charge carriers. 1–12 (2007). <https://doi.org/10.1103/PhysRevB.76.045210>
- Miller, A. & Abrahams, E. Impurity conduction at low concentrations. *Phys. Rev.* **120**, 745–755 (1960).
- Udrescu, S. M. & Tegmark, M. AI Feynman: A physics-inspired method for symbolic regression. *Sci. Adv.* **6**, eaay2631 (2020).
- Udrescu, S. M. *et al.* AI Feynman 2.0: Pareto-optimal symbolic regression exploiting graph modularity. *Adv. Neural Inf. Process. Syst.* **33**, 1–12 (2020).
- Long, Z., Lu, Y. & Dong, B. PDE-Net 20: Learning PDEs from data with a numeric-symbolic hybrid deep network. *J. Comput. Phys.* **399**, 108925 (2019).
- Raissi, M., Yazdani, A. & Karniadakis, G. E. Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations. *Science (80-)*. **367**, 1026–1030 (2020).
- Raissi, M., Perdikaris, P. & Karniadakis, G. E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **378**, 686–707 (2019).
- Coehoorn, R., Pasveer, W. F., Bobbert, P. A. & Michels, M. A. J. Charge-carrier concentration dependence of the hopping mobility in organic materials with. 1–20 (2005). <https://doi.org/10.1103/PhysRevB.72.155206>
- Fishchuk, I. I. *et al.* Temperature dependence of the charge carrier mobility in disordered organic semiconductors at large carrier concentrations. 1–12 (2010). <https://doi.org/10.1103/PhysRevB.81.045202>
- Schmechel, R. Gaussian disorder model for high carrier densities: Theoretical aspects and application to experiments. *Phys. Rev. B* **66**, 235206 (2002).
- Klauk, H. Organic thin-film transistors. *Chem. Soc. Rev.* **39**, 2643 (2010).
- Meijer, E. J., Matters, M., Herwig, P. T., de Leeuw, D. M. & Klapwijk, T. M. The Meyer-Neldel rule in organic thin-film transistors. *Appl. Phys. Lett.* **76**, 3433 (2000).
- Meijer, E. Ph.D. Thesis: Charge transport in disordered organic field-effect transistors. (Technical University of Delft, 2003).
- Ullah, M. *et al.* Dependence of Meyer-Neldel energy on energetic disorder in organic field effect transistors. *Appl. Phys. Lett.* **96**, 1–4 (2010).

25. Sitter, H., Stifter, D. & Nguyen Manh, T. Preparation of pristine and Ba-doped C60 films by hot-wall epitaxy. *J. Cryst. Growth* **174**, 828–836 (1997).
26. GitHub - trevorstephens/gplearn at ad57cb18caafdb02cca861aea712f1bf3ed5016e. <https://github.com/trevorstephens/gplearn/tree/ad57cb18caafdb02cca861aea712f1bf3ed5016e>. (Accessed 7th February 2021)
27. Wang, Y., Wagner, N. & Rondinelli, J. M. Symbolic regression in materials science. *MRS Commun.* **9**, 793–805 (2019).
28. Sun, S., Ouyang, R., Zhang, B. & Zhang, T. Y. Data-driven discovery of formulas by symbolic regression. *MRS Bull.* **44**, 559–564 (2019).
29. Virgolin, M., Alderliesten, T., Bel, A., Witteveen, C. & Bosman, P. A. N. Symbolic regression and feature construction with GP-GOMEA applied to radiotherapy dose reconstruction of childhood cancer survivors. in *GECCO 2018—Proceedings of the 2018 Genetic and Evolutionary Computation Conference* 1395–1402 (Association for Computing Machinery, Inc, 2018). <https://doi.org/10.1145/3205455.3205604>
30. Franco, V., Conde, A., Romero-Enrique, J. M. & Blázquez, J. S. A universal curve for the magnetocaloric effect: An analysis based on scaling relations. *J. Phys. Condens. Matter* **20**, 285207 (2008).
31. Repaka, D. V. M., Tripathi, T. S., Aparnadevi, M. & Mahendiran, R. Magnetocaloric effect and magnetothermopower in the room temperature ferromagnet  $\text{Pr}_{0.6}\text{Sr}_{0.4}\text{MnO}_3$ . *J. Appl. Phys.* **112**, 123915 (2012).
32. Pasveer, W. F. *et al.* Unified description of charge-carrier mobilities in disordered semiconducting polymers. *Phys. Rev. Lett.* **94**, 206601 (2005).
33. Cottaar, J., Koster, L. J. A., Coehoorn, R. & Bobbert, P. A. Scaling theory for percolative charge transport in disordered molecular semiconductors. *Phys. Rev. Lett.* **107**, 136601 (2011).
34. Chang, J. F., Siringhaus, H., Giles, M., Heeney, M. & McCulloch, I. Relative importance of polaron activation and disorder on charge transport in high-mobility conjugated polymer field-effect transistors. *Phys. Rev. B Condens. Matter Mater. Phys.* **76**, 205204 (2007).
35. Chang, J. F. *et al.* Enhanced Mobility of poly(3-hexylthiophene) transistors by spin-coating from high-boiling-point solvents. *Chem. Mater.* **16**, 4772 (2004).
36. Zhang, Y. *et al.* Probing carrier transport and structure-property relationship of highly ordered organic semiconductors at the two-dimensional limit. *Phys. Rev. Lett.* **116**, 016602 (2016).

## Acknowledgements

R.D., S.J. and K.H. acknowledge funding from the Accelerated Materials Development for Manufacturing Program at A\*STAR via the AME Programmatic Fund by the Agency for Science, Technology and Research under Grant No. A1898b0043. W.L.L. would like to acknowledge funding support from Ministry of Education (MOE) under AcRF Tier 2 Grant (2019-T2-2-106).

## Author contributions

M.L., R.D., D.V.M.R., and K.H. developed the concept presented in this paper. M.L. and R.D. carried out the computation and drafted the manuscript. K.H., W.L.L., and S.J. supervised the entire project. All authors read and approved the final manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-02737-7>.

**Correspondence** and requests for materials should be addressed to S.J., W.L.L. or K.H.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021