Electronic Supplementary Material (ESI) for Energy & Environmental Science.

Supplementary Information for

Predicting the photocurrent-composition dependence in organic solar cells

Xabier Rodríguez-Martínez[a], Enrique Pascual-San-José[a], Zhuping Fei[b], Martin Heeney[b], Roger Guimerà[c,d] and Mariano Campoy-Quiles*[a]

[a] Institut de Ciència de Materials de Barcelona, ICMAB-CSIC, Campus UAB, 08193, Bellaterra, Spain.

[b] Department of Chemistry and Centre for Plastic Electronics, White City Campus, Imperial College London, London W12 0BZ, United Kingdom.

[c] Institució Catalana de Recerca i Estudis Avançats, ICREA, Passeig de Lluís Companys 23, 08010, Barcelona, Spain.

[d] Department of Chemical Engineering, Universitat Rovira i Virgili, 43007, Tarragona, Spain.

* Author to whom correspondence should be addressed: m.campoy@csic.es

Section I. Materials and methods

a. Device fabrication

All devices were manufactured with inverted architecture (glass/ITO/ZnO/active layer/MoO_3/Ag). Pre-patterned and non-patterned ITO substrates (purchased from Ossila, 100 nm thick and 20 Ω square^{-1} sheet resistance), were cleaned by sequential ultrasonication in acetone, Hellmanex 10% solution in water, isopropanol and finally 10% NaOH water solution. A dispersion of ZnO nanoparticles (N-10, Avantama) was blade coated onto cleaned ITO substrates in air conditions using an automatic coater (ZAA 2300, Zehntner) and a lamination piece of equipment (ZUA 2000, Zehntner). The ZnO casting parameters were 50 μm blade gap, 50 μL casting volume, 4 mm s^{-1} blade speed and 40 °C yielding a thickness of 40 nm. A 100 °C post-deposition thermal annealing was carried out for 10 min in air. The ZnO coated substrates were then transferred into a nitrogen-filled glovebox for the deposition of the active layers.

All photovoltaic materials were used as received and dissolved in chlorobenzene with a total solid concentration of 15 mg mL^{-1} and stirred overnight at 80 °C. PBT7-Th and PCDTBT were purchased from Ossila. PBDB-T, PBDB-T-2Cl, PBDB-T-2F, ITIC, ITIC-M and ITIC-4F were purchased from 1-Material. ITIC-C_8 and ITIC-C_2C_6 were synthesized following previous publications of one of our groups and are further detailed below[1]. $PC_{70}BM$ was purchased from Solenne. On the one hand, discrete 1D-thickness graded samples (Figure S1a) were manufactured using premixed inks of donor and acceptor materials weighted in controlled fractions in a single vial, resulting in devices with fixed composition. D:A mixing ratio solutions extending from 1:0 to 0:1 were individually prepared to screen the composition space for PTB7-Th:ITIC and PBDB-T:ITIC-C_2C_6. The casting conditions were 100 µm blade gap, 70 µL casting volume and 90 °C casting temperature. The blade was intendedly decelerated from 90 mm s^{-1} to 5 mm s^{-1} resulting in a thickness-graded active layer (ca. 50-200 nm)[2,3]. On the other hand, 2D-thickness-composition graded samples (Figure S1b) were manufactured following the same processing parameters as 1D devices excepting the blade speed, which in this case was kept fixed at 5 mm s^{-1} and two drops of 40 µL each of the neat material inks were cast. Finally, 10 nm of MoO_3 and 100 nm of Ag were thermally evaporated at a rate of 0.1 and 1 Å s^{-1}, respectively. Note that each 1D sample contains 24 devices, 12 of them with different thickness values (two per side), with a pixel active area of 8 mm^2.

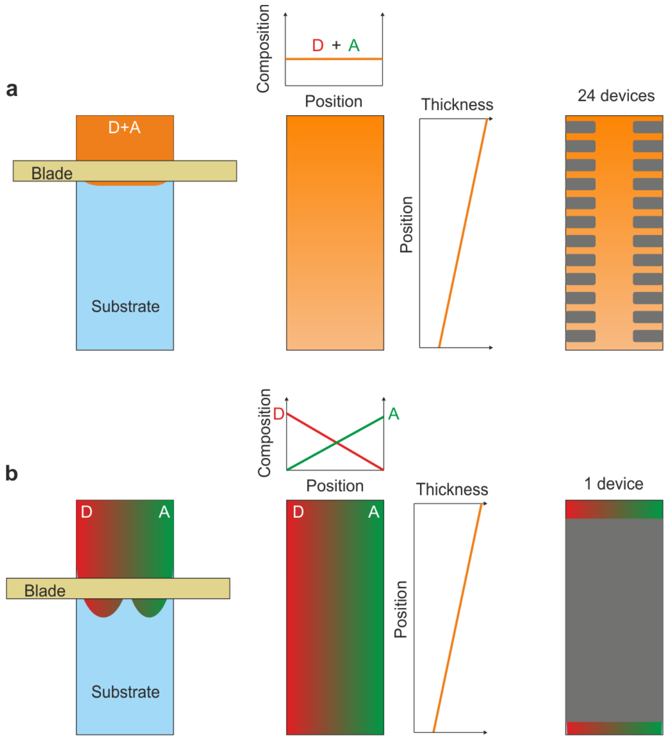


Figure S1. The two processing schemes for organic solar cells comprising lateral gradients in their active layer. (a) By decelerating the applicator during the coating of the active layer we generate a thickness gradient for inks of fixed D:A ratio. This approach, in combination with pre-patterned ITO substrates and their dedicated evaporation masks, allows the high-throughput screening of the thickness-performance dependence in 24 discrete devices. (b) To simultaneously screen very efficiently the composition, two drops of the pristine D and A inks coalesce during the coating leading to an orthogonal distribution of D:A ratio and active layer thickness. In this case, a single large-area device is obtained, which is then mapped by LBIC and Raman spectroscopy imaging.

b. Device characterization

J-V characteristics of 1D devices were automatically extracted by using a Keithley source meter and an Arduino based multiplexer/switcher that allows measuring 24 devices in less than 6 minutes. As a lighting source, a SAN-EI Electric, XES-100S1 AAA solar simulator was used to ensure a homogenous illumination in a 10 cm x 10 cm area. The solar simulator was previously calibrated with a certified silicon solar cell (NREL).

2D gradient large-area devices were characterized using a WITec alpha 300 RA+ confocal setup, connected to a current amplifier to extract photocurrent maps of the solar cell. The samples were excited using a 10X objective (NA 0.25) coupled to a 488 nm solid-state laser (with its power set to 3 mW to avoid photodegradation). Our optical configuration yields a ca. 10 μm laser spot in diameter once focused on the active layer. Light beam induced current (LBIC) with monochromatic light was measured simultaneously along with Raman in short-circuit conditions. However, to extend the electromagnetic spectrum of the excitation source, we coupled the built-in lamp of the WITec alpha 300 RA+ piece of equipment and employed it to map the photovoltaic response under white light illumination (Figure S2). Notably, the matching of the lamp irradiance with the solar spectrum is limited. As a result, we might expect photocurrent underestimations when employing high band gap (<500 nm, UV) or low band gap (>700 nm, IR) materials, which are those employed more frequently in semi-transparent OPV devices. The use of white light yielded a spot size of ca. 150 μm in diameter. Data analysis (D:A mixing ratio and thickness from Raman spectra) was performed using homemade MATLAB routines described elsewhere[4].

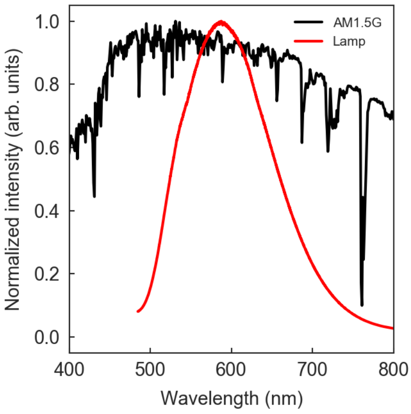
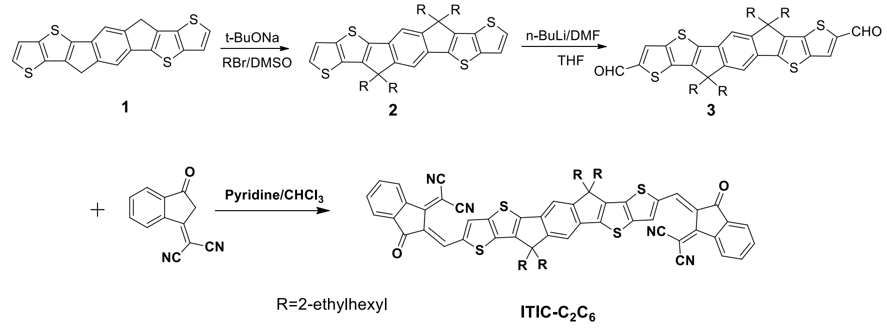


Figure S2. Normalized emission spectrum of the built-in lamp in the WITec alpha 300 RA+ setup measured upon reflection on a silver mirror (solid red line). For comparative purposes, the reference AM1.5G irradiance spectrum is also shown (solid black line) to indicate the notorious mismatch in regions far from the visible region of the electromagnetic spectrum.

c. Synthesis and characterization of ITIC-C_2C_6



Scheme S1. The synthetic route of ITIC-C_2C_6.

5,11-Dihydro-5,5,11,11-(2-ethylhexyl)-*s*-indaceno[1,2-*b*:5,6-*b*'] dithieno[3,2-b]thiophene (2):

To a suspension of 5,11-dihydro-*s*-indaceno[1,2-*b*:5,6-*b*'] dithieno[3,2-*b*]thiophene (**1**) (0.70 g, 1.85 mmol) in anhydrous DMSO (40 ml) was added sodium *tert*-butoxide (1.06 g, 11.11 mmol) in parts. The reaction mixture was heated at 80 °C for 1 h, followed by the addition of 1-bromohexadecane (2.14 g, 11.11 mmol) dropwise. After complete addition, the resultant mixture was heated at 80 °C for overnight, then poured into ice-water and extracted by THF/hex (2/1, v/v, 100 mL) for 3 times. The combined organics were dried by MgSO4, filtered and concentrated under reduced pressure. This residue was purified by column chromatography on silica (eluent: hexane) to give a yellow oil (0.72 g, 47%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.35 (d, 2H), 7.34-7.29 (m, 4H), 2.11-2.06 (m, 8H), 0.99-0.73(m, 32H), 0.70-0.45 (m, 28H); ^{13}C NMR (100 MHz, $CDCl_3$): δ 152.2, 145.4, 143.8, 141.0, 136.6, 134.2, 125.5, 120.3, 113.9, 53.8, 35.0, 34.9, 33.7, 29.1, 28.5, 28.4, 28.0, 27.0, 26.8, 22.9, 22.8, 22.7, 14.2, 13.8, 10.7, 10.6, 10.2.

6,6,12,12-Tetra(2-ethylhexyl)-6,12-dihydrothieno[3,2-b]thieno[2'',3'':4',5']thieno[2',3':5,6]-s-indaceno[2,1-d]thiophene-2,8-dicarbaldehyde (3):

A solution of n-BuLi (2.1 mL of a 1.6 M solution in pentane, 3.36 mmol) was added dropwise into **2** (0.70 g, 0.85 mmol) in THF (40 mL) at -78 °C. After stirred at this temperature for 20 min, the mixture was allowed to warm to RT for 1.5 h. Then the mixture was cooled to -78 °C , and anhydrous DMF (0.3 mL) was added in one portion, and the mixture was allowed to warm to RT slowly and stirred overnight. Water (100 mL) was added, and the mixture was extracted with DCM (3 × 200 mL). The combined organics were dried by $MgSO_4$, filtered and concentrated under reduced pressure. The residue was purified by silica gel chromatography (eluent: DCM/hexane = 3/2, v/v) to afford an orange oil (0.50 g, yield: 67%). ^{1}H NMR ($CDCl_3$, 400 MHz,) δ (ppm): 9.96 (s, 2H), 8.00 (s, 2H), 7.48 (s, 2H), 2.16-2.10 (m, 8H), 0.99-0.73 (m, 32H), 0.70-0.44 (m, 28H); ^{13}C NMR ($CDCl_3$, 100 MHz,) δ (ppm): 182.9, 153.7, 150.4, 146.2, 143.9, 141.3, 140.8, 137.0, 130.2, 115.1, 54.2, 54.1, 43.2, 43.1, 42.6, 35.3, 35.1, 35.0, 33.8, 33.4, 28.5, 28.4, 28.0, 27.1, 27.0, 2687, 26.7, 22.8, 22.7, 22.6, 14.2, 13.8, 10.8, 10.6, 10.1.

{(2Z)-2-[(8-{(E)-[1-(Dicyanomethylidene)-3-oxo-1,3-dihydro-2H-inden-2-ylidene]methyl}-6,6,12,12-tetra(2-ethylhexyl)-6,12-dihydrothieno[3,2-b]thieno[2'',3'':4',5']thieno[2',3':5,6]-s-indaceno[2,1-d]thiophen-2-yl)methylidene]-3-oxo-2,3-dihydro-1H-inden-1-ylidene}propanedinitrile (ITIC-C_2C_6):

To a mixture of 3-dihydro-1H-inden-1-ylidene)malononitrile (165 mg, 0.85 mmol) and compound **3** (150 mg, 0.17 mmol) in $CHCl_3$ (20 mL) was added pyridine (0.5 mL). After addition, the mixture was heated to 80 °C and stirred for overnight, and then poured into water (30 mL) and extracted with DCM (3 × 20 mL). The combined organics were dried by $MgSO_4$, filtered and concentrated under reduced pressure. The residue was purified by silica gel chromatography (eluent: DCM/hexane = 2/1), followed by recrystallization from DCM/MeOH to afford a shining golden solid (143 g, yield: 68%). ^{1}H NMR ($CDCl_3$, 400 MHz,) δ (ppm): 8.99 (s, 2H), 8.71 (d, 2H), 8.21 (s, 2H), 7.96 (d, 2H), 7.79-7.76 (m, 4H), 7.54 (d, 2H), 2.29-2.16 (m, 8H), 0.96-0.71 (40H), 0.70-0.46 (m, 20H); ^{13}C NMR ($CDCl_3$, 100 MHz,) δ (ppm): ^{13}C NMR (100 MHz, $CDCl_3$) δ (ppm): 188.3, 160.5, 154.8, 153.7, 147.9, 147.4, 143.2, 140.1, 139.3, 138.4, 137.7, 137.3, 136.9, 135.2, 134.5, 125.3, 123.8, 122.4, 115.6, 114.7, 69.2, 54.3, 43.1, 43.0, 42.8, 35.4, 35.2, 33.7, 33.5, 28.6, 28.4, 28.1, 27.2, 27.0, 26.8, 22.8, 22.7, 22.6, 14.2, 13.8, 10.8, 10.2.

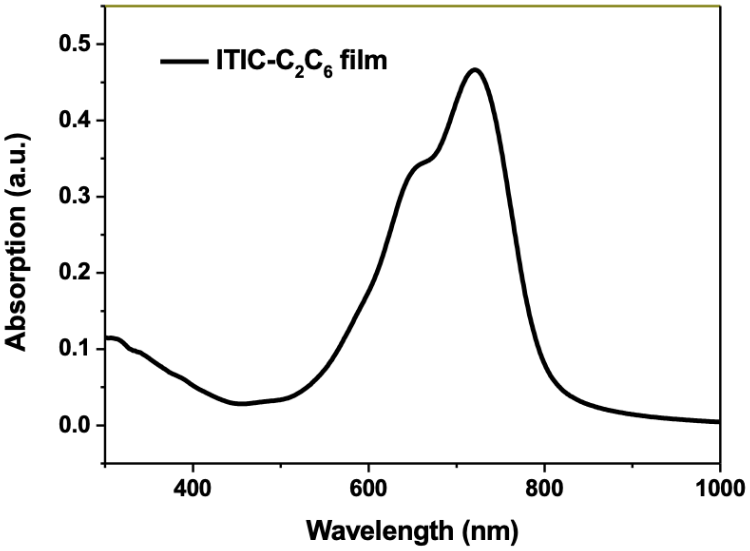


Figure S3. UV-Vis absorption spectrum of an ITIC-C_2C_6 film. The main absorption peaks locate at 656 nm and 721 nm. The absorption edge locates at 800 nm, corresponding to an optical band gap of 1.55 eV. PESA data for ITIC-C_2C_6 shows an ionization potential of 5.71 eV, which compared to ITIC-C_8 (IP = 5.68 eV from PESA and HOMO = -5.63 eV from CV) allows setting the HOMO level of ITIC-C_2C_6 at -5.66 eV.

Section II. Comparison of Edisonian and high-throughput experimentation approaches

a. Sampling the thickness-composition space by means of active layer thickness gradients

For comparative purposes against the high-throughput screening methodology results, we prepared multiple solar cell devices in which the D:A blend ratio was varied in discrete samples and controlled steps throughout the entire compositional range, including in some cases neat active layers (i.e. homojunctions). This is the classical approach for D:A blend ratio screening, also known as Edisonian experimentation: the preparation of individual devices, each from a compositionally homogenized ink that results from the overnight stirring of controlled D:A weights in a single vial. However, to accelerate the corresponding active layer thickness screening, the devices were blade coated to include an active layer thickness gradient spanning from ca. 50 nm to 200 nm for each composition step. This approach, combined with pre-patterned 25x75 mm^2 ITO substrates, enabled us to screen up to 12 different active layer thickness values for each D:A composition step, thus speeding up and making more efficient (in time and resources) the screening process.

Figure S4 depicts the solar cell device parameters (J_{sc}, V_{oc}, FF and PCE) obtained in PTB7-Th:ITIC (upper row) and PBDB-T:ITIC-C_2C_6 (lower row) batches, which are those included in Fig. 2 in the main manuscript. We first observe that V_{oc} is fairly constant throughout the entire D:A composition diagram. Conversely, the FF varies more abruptly with the composition, showing its maximum close to 20-30 wt% of donor in PTB7-Th:ITIC blends and 35-40 wt% of donor in PBDB-T:ITIC-C_2C_6 blends. Nevertheless, J_{sc} is still a good probe for PCE as they both show very similar upper-shell data distributions (green dashed lines) and peak performance positions in the D:A composition space.

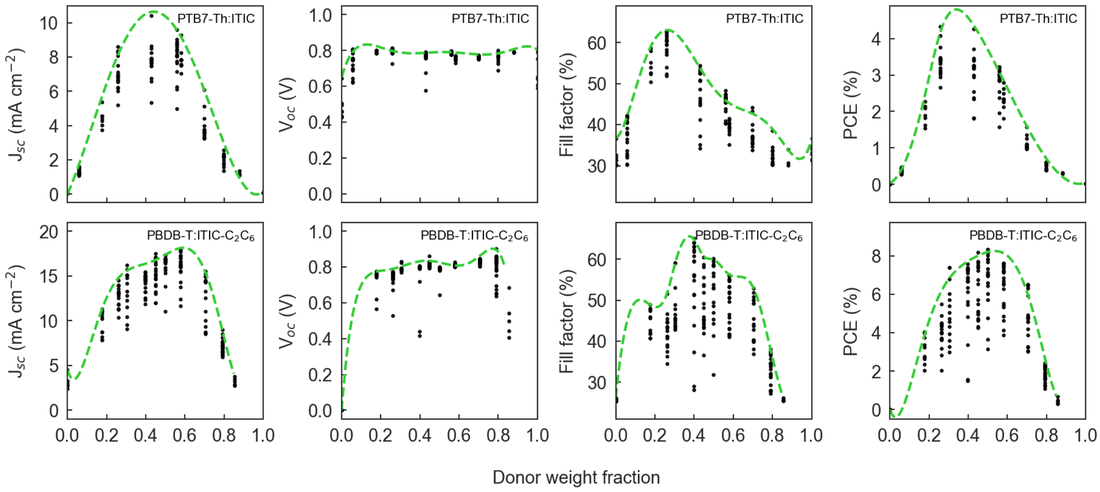


Figure S4. Solar cell device parameters (J_{sc}, V_{oc}, FF and PCE) for discrete composition step batches of PTB7-Th:ITIC (upper row) and PBDB-T:ITIC-C_2C_6 (lower row). Since the devices were processed to show an active layer thickness gradient to accelerate its corresponding screening, the observed dispersion in the *y*-axis is exclusively ascribed to such deliberate thickness variation.

b. Analysis of the time and material cost savings in high-throughput experimentation

Despite the significant savings in time and resources offered by the (1D) thickness graded devices compared to the traditional discrete sampling approach, the use of bidimensional (2D) thickness-composition graded devices extends even further the abovementioned savings. A detailed time and material cost analysis of the corresponding discrete and 1D/2D graded devices is listed in Table S1. According to our estimations and compared to 1D graded devices, 2D combinatorial devices offer a 2,000-fold higher density of data points per sample (12 discrete pixels per 1D sample vs 24,000 data points per 2D sample) as well as an eleven-fold reduction in the usage of raw semiconducting materials to perform the J_{sc}-vol% dependence exploration. As a result, by following the 2D high-throughput screening approach the experimental time required per data point is below 90 seconds and the consumption of raw semiconducting material is as low as a few tens of nanograms per data point (Table S1).

Table S1. Time and material cost analysis for discrete sampling, 1D thickness-graded devices and 2D thickness-composition graded devices (high-throughput experimentation) in the full exploration of the J_{sc}-vol% space for a given OPV binary.

	Discrete sampling	1D thickness gradient	2D thickness-composition gradients
No. samples	66 [a]	11 [b]	1
No. data points per sample	1	12	24,000
No. data points (total)	66	132	24,000
Manufacturing time (h)	93.2 [c]	16.5	5.9
Measuring time (h)	0.55 [d]	1.12 [c]	9
Data analysis (h)	1	2	13
Total time (h)	94.8	19.6	27.9
mg of semiconductor	79.2	13.2	1.2
Time required per data point (h)	1.44	0.15	0.0012 (86 s)
mg required per data point	1.2	0.1	5E-05 (50 ng)

[a] From, to neat films in 10 vol% steps + 6 homogeneous thickness steps

[b] From, to neat films in 10 vol% steps + 12 thickness steps as gradient

[c] Assuming 11 batches with co-evaporation of 6 samples per batch

[d] J-V curve takes 30 s to be measured

c. Reproducibility evaluation of the high-throughput experimentation approach

We tested the reproducibility of the developed high-throughput experimentation approach by fabricating two batches of composition graded devices for the same D:A pair (PBDB-T:ITIC-C_8) in different days. Their corresponding large area maps of active layer thickness, donor volume fraction and normalized photocurrent are depicted in Figure S5. The spatial correlation of all three magnitudes in both batches is accordingly illustrated in Figure S6 and it demonstrates that the high-throughput experimentation approach is parametrically highly reproducible and self-consistent in terms of normalized photocurrent.

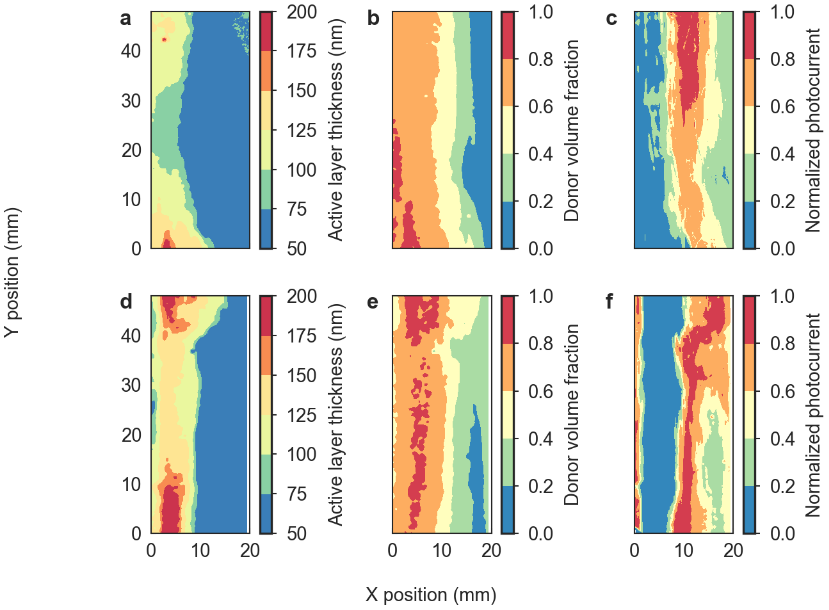


Figure S5. Batch-to-batch comparison of PBDB-T:ITIC-C_8 combinatorial devices. Active layer thickness maps (a,d) and composition images (b,e) were acquired by properly deconvoluting the spatially-resolved Raman spectra.[4] The normalized photocurrent images (c,f) were acquired in a subsequent step by mapping of a 150 µm white light spot throughout the active area of the device in the same Raman imaging setup.

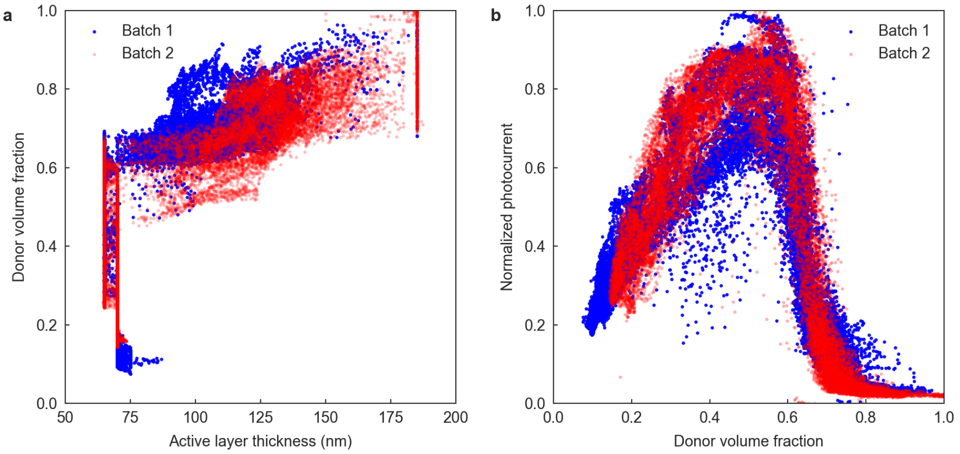


Figure S6. The one-to-one comparison of the PBDB-T:ITIC-C_8 batches demonstrates that (a) a similar region of the composition vs. active layer thickness diagram is accessed, and (b) that in terms of normalized photocurrent the overall trend is overlapping in both cases.

d. Morphological characterization of the active layers

Verifying that *in situ* mixing during blade coating leads to the same degree of mixing attained using fully premixed D:A inks can further support the robustness of the conclusions drawn from combinatorial devices. In that regard, we have recently shown that Raman shifts and photoluminescence (PL quenching and shifts) can be correlated to microstructure in high-throughput experiments.[5,6] Raman has, indeed, been used in the literature to address the degree of aggregation in bulk heterojunctions.[7] Figure S7a depicts the Raman shift of the main peaks corresponding to either PBDB-T or ITIC-C_2C_6 (with respect to the corresponding pristine materials, i.e. $\Delta\omega = \omega_{\text{blend}} - \omega_{\text{pristine}}$) as a function of composition for both, discrete pixels (same as the devices in Figure S4) and compositional gradients (Figure 2a,b). As a rule of thumb, the larger the difference between the Raman peak and the initial vibration energy (i.e. further blue-shifted spectra), the larger the disruption of the local environment. We observe:

1. As one material is more diluted, there is a stronger variation of the corresponding Raman spectra. This is indicative of some degree of miscibility between materials.
2. Judging by the Raman shifts, the gradient samples show different degrees of mixing for any given composition, including the best mixing possible, i.e. that obtained when the solutions are premixed and here represented by the single data points from devices contained in Figure S4. The fact that these well-mixed data points lie at the higher extreme (in terms of Raman shift) of the gradient samples suggests that the other attained microstructures in that sample are more phase separated than when the materials are premixed in solution. This also probes that the catalogue of film morphologies explored in compositional gradients is richer than in conventional devices.
3. The Raman shift difference between donor and acceptor (Figure S7b) indicates that the maximum J_{sc} occurs at values close to 3-3.5 cm^{-1}, which might be indicative of an optimized degree of phase separation between the blended species. Interestingly, we observe that compositional gradients (cloud of scattered points) reproduce very well the overall photocurrent trend observed in discrete devices (solid line), but providing much more statistical relevance.

In addition to Raman shifts we have evaluated further potential priors about morphology, such as the PL. For this, we have measured PL maps at 633 nm excitation in our PBDB-T:ITIC-C_2C_6 graded device (so that both materials are absorbing), in the same setup used to extract Raman. From the scattered data shown in Figure S8 we inferred the following trends in terms of PL intensity and shifting:

1. Photocurrent is optimized at submaximal PL quenching scenarios. This finding supports that a certain degree of phase separation is required to maximize simultaneously charge photogeneration and charge extraction, probably due to a detrimental charge carrier mobility in one of the species (likely to be the donor as suggested by our literature revision in Figure S14 and the large Raman shifts upon blending).

2. Photocurrent achieves its maximum when the PL spectrum is centred in the range of 780-800 nm, which could be indicative of enhanced crystallinity in the acceptor phase (as showing the more red-shifted PL emission).

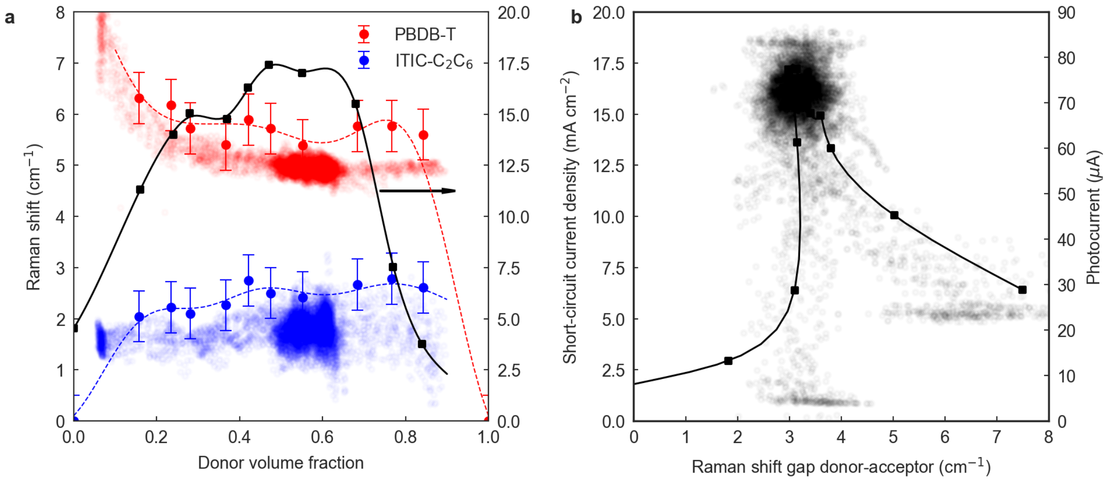


Figure S7. (a) Raman shift of the main vibrational fingerprints of PBDB-T (red circles) and ITIC-C$_2$C$_6$ (blue circles) obtained in discrete devices following traditional experimentation workflows (solid circles with error bars) and graded devices (depicted as semi-transparent clouds of points). The solid black line corresponds to a guide-to-the-eye of the short-circuit current density (J_{sc}) measured in the best performing discrete devices in terms of PCE. (b) A correlation plot of J_{sc} and photocurrent vs Raman shift difference between donor and acceptor indicates that the morphology for optimum photogeneration is attained at Raman shift gaps close to 3-3.5 cm^{-1}.

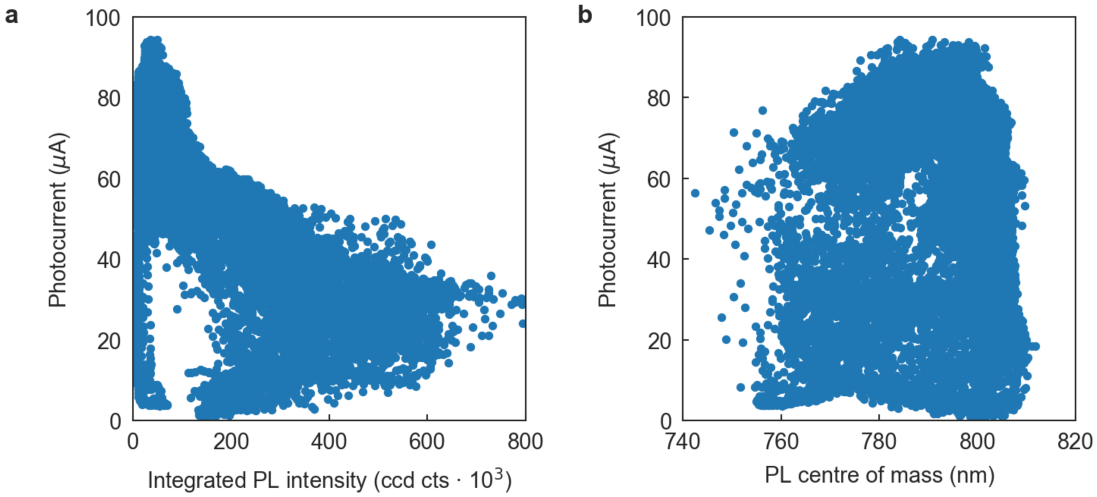


Figure S8. (a) Photocurrent vs. integrated PL intensity shows that submaximal PL quenching scenarios correlate well with enhanced photogeneration and charge extraction. (b) Photocurrent vs. PL emission centre shows a broader distribution of values although we observe that photocurrent is precisely enhanced when the PL is centred close to 780-800 nm.

e. Relationships between Raman, PL, active layer thickness and FF in discrete and combinatorial devices

Since focusing solely on J_{sc} when optimizing D:A OPV blends might lead in some cases to a suboptimal PCE (mainly due to a dissimilar FF dependence with D:A ratio, see Figure S4), we have evaluated potential training features that could mitigate this effect by modelling also the FF dependence on composition. FF depends on the total active layer thickness, the microstructure, charge transport properties (including the existence of a space charge limited current (SCLC) scenario), the quality of the interlayers and even the active area of the devices, all having important effects on the series and shunt resistances.[8] To extend the AI study to investigate FF we would need initial input parameters that characterize the aforementioned dependencies and, moreover, much larger data statistics, as we only made discrete devices for two binaries in the present work (those shown in Figure S4). As a preliminary investigation, we have evaluated if Raman, PL or thickness can be used as prior for the FF dependence on D:A ratio.

First, we explored the possibility for the Raman shift of the main vibrational modes of the blended materials to be a potential prior of the FF variations (Figure S9). In this case, we found that the FF peaks close to a D-A Raman shift gap ($\Delta\omega_{donor}$ - $\Delta\omega_{acceptor}$) of 3.5 cm^{-1}, indicating the formation of a slightly distinct morphology as for maximum photocurrent (ca. 3 cm^{-1}). In this respect, electronic transport might be the key element to determine FF for a given device architecture.

The PL intensity and shift could also be included as potential parameters correlated to FF. As illustrated in Figure S10a, in PBDB-T:ITIC-C_2C_6 blends the FF maximizes at intermediate PL quenching scenarios. This finding suggests that photogeneration is the dominant factor in the compositional window defined by the main photocurrent maximum (50-60 vol% of donor) whereas the maximum FF occurs where some degree of phase separation is present, likely to improve charge transport. Furthermore, an analysis of the PL shifting with FF (Figure S10b) indicates that largely red-shifted PL distributions (ca. 800 nm) might correlate better with an enhanced FF in the compositional window of 30-40 vol% of donor. This could be indicative of the formation of ITIC-C_2C_6 aggregates as these are responsible for red-shifted PL. This is in good agreement with our previous observation regarding PL quenching and photocurrent as ITIC-C_2C_6 shows an improved charge transport compared to PBDB-T (Figure S14).

Finally, we have also observed that at intermediate D:A ratios the FF correlates reasonably well with thicker active layers, as suggested by the trends found as a function of the pixel number in discrete devices (Figure S11). With these encouraging results in mind, we conclude that Raman shifting, PL quenching and PL shifting (including active layer thickness) could be incorporated as additional input variables in upcoming AI models to better account for morphology, including its ultimate relationship with J_{sc}, FF and eventually PCE.

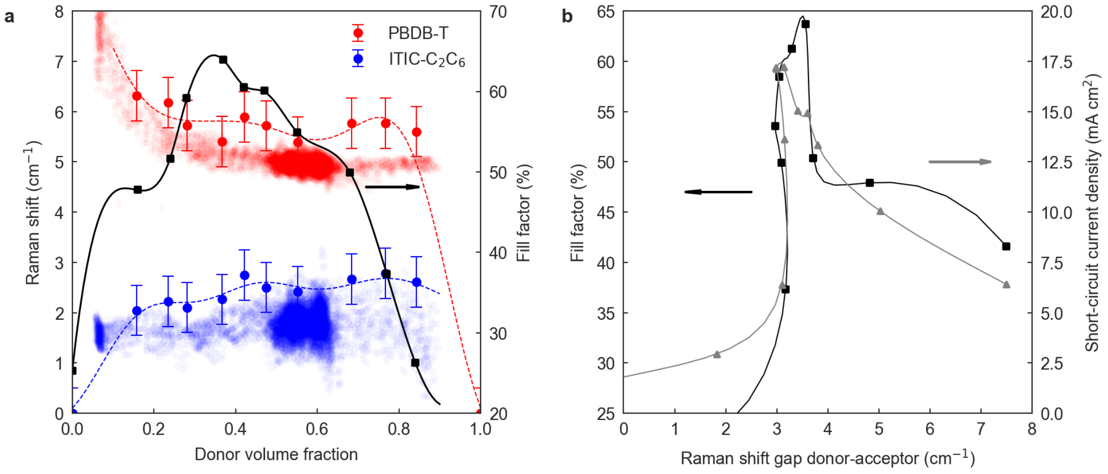


Figure S9. (a) Raman shift of the main vibrational fingerprints of PBDB-T (red circles) and ITIC-C_2C_6 (blue circles) obtained in discrete devices following traditional experimentation workflows (solid circles with error bars) and graded devices (depicted as semi-transparent clouds of points). The solid black line corresponds to a guide-to-the-eye of the FF measured in the best performing discrete devices in terms of PCE. (b) A correlation plot of FF and photocurrent vs Raman shift difference between donor and acceptor indicates that the morphology for optimum photogeneration is attained at Raman shift gaps close to 3 cm^{-1}, while the FF peaks at slightly higher shift gaps (3.5 cm^{-1}).

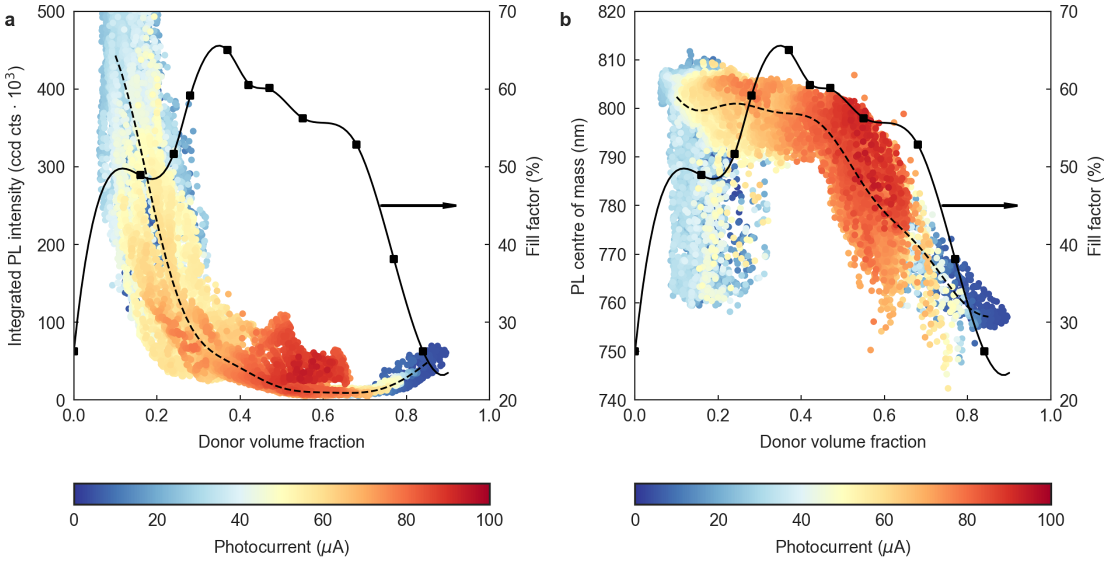


Figure S10. PL data for the PBDB-T:ITIC-C_2C_6 graded device extracted at 633 nm excitation. (a) The integrated PL intensity distribution shows its minimum (i.e. maximum quenching) at the compositional window where photocurrent attains its absolute maximum. This is indicative of photogeneration being the dominant factor at such compositional regime. (b) The PL shifting distribution indicates that the maximum FF is achieved for red-shifted PL spectra compared to those collected at the maximum in photocurrent. Dashed lines represent (in both panels) support vector machine (SVM) models to the experimental data to draw the main statistical trendlines. Solid lines are splines connecting the data observed in the discrete devices.

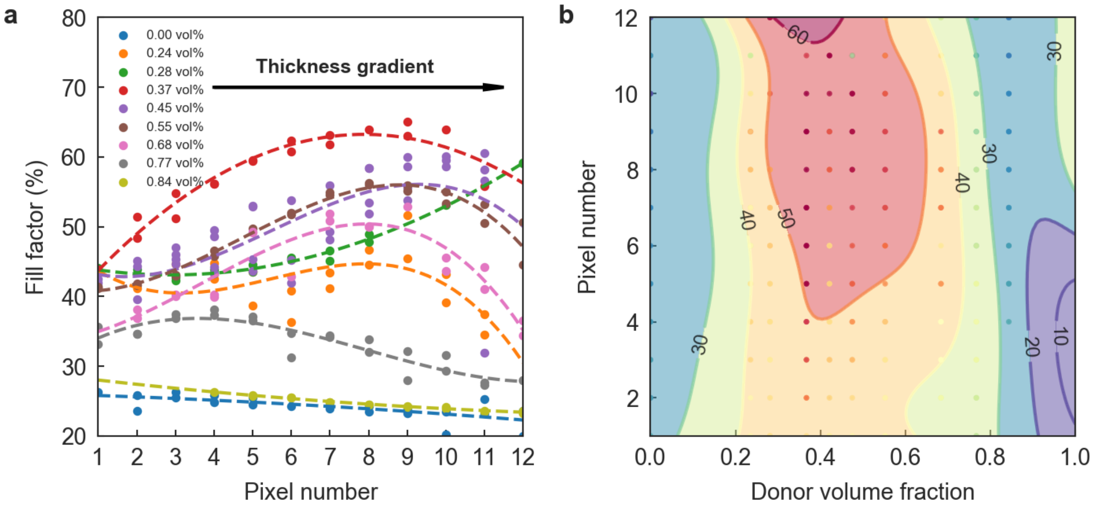


Figure S11. (a) The FF dependence as a function of the pixel number in discrete PBDB-T:ITIC-C_2C_6 devices (containing 1D thickness gradients) shows that at balanced D:A ratios the active layer thickness largely determines the FF. (b) A SVM model regressor is applied to our experimental data to obtain a smoother landscape relating pixel number (i.e. thickness), composition and FF (in colour scale).

Section III. Ellipsometric characterization: normalized absorption coefficients and Tauc plots

The absorption coefficients (α) were deduced from the corresponding extinction coefficients (κ), extracted by means of variable angle spectroscopic ellipsometry (VASE) using a Semilab GES5E rotating polarizer ellipsometer. Modeling of the ellipsometry data was performed using the Winelli II piece of software package from SOPRALAB.

For the determination of the solid-state optical band gaps using the Tauc plots, we depict the magnitude $(\alpha\, h\nu)^{1/r}$ vs $h\nu$ (photon energy in eV), where $r = 2$ for indirect allowed transitions. By consistently extrapolating the linear regime of the absorption onset we determine the corresponding optical band gaps.

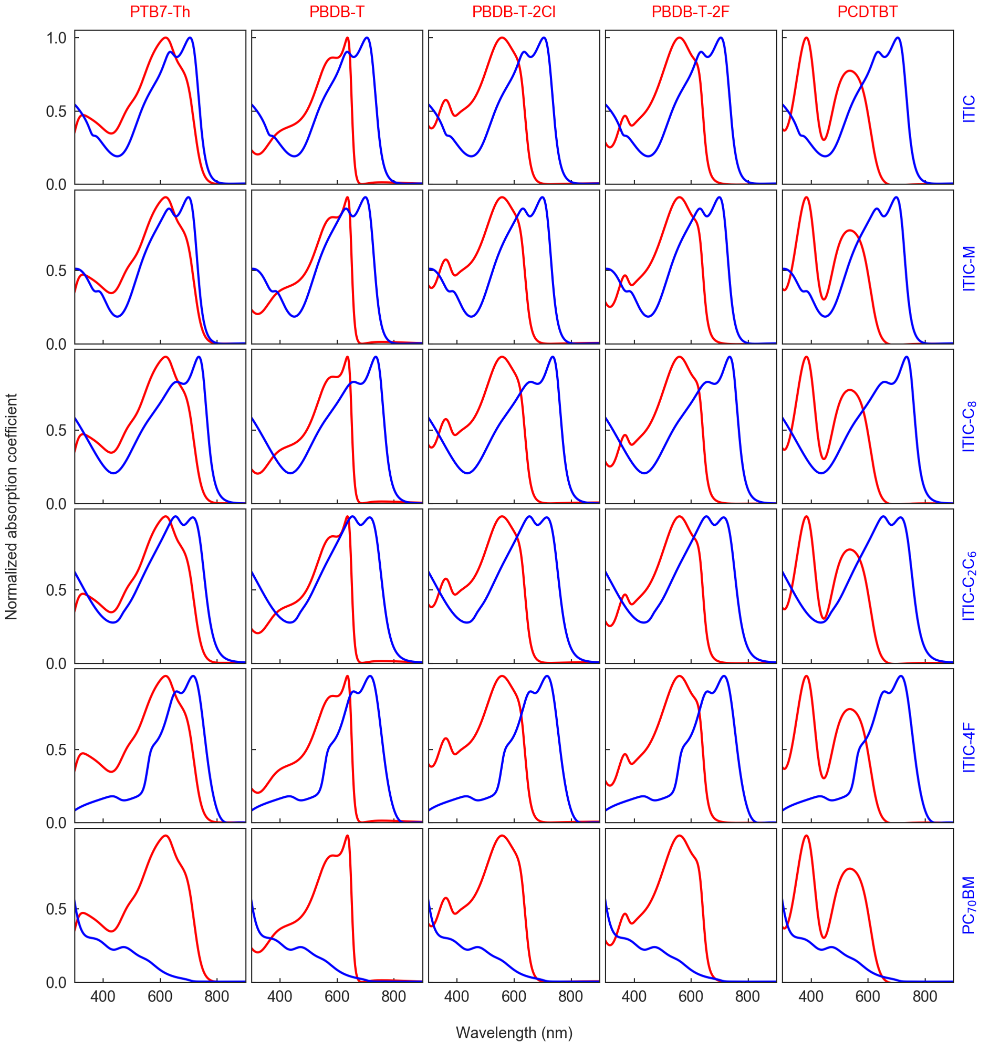


Figure S12. Combinatorial matrix of normalized absorption coefficients corresponding to the D:A binary OPV systems screened in the main manuscript. Data for donors and acceptors are depicted in red and blue colors, respectively.

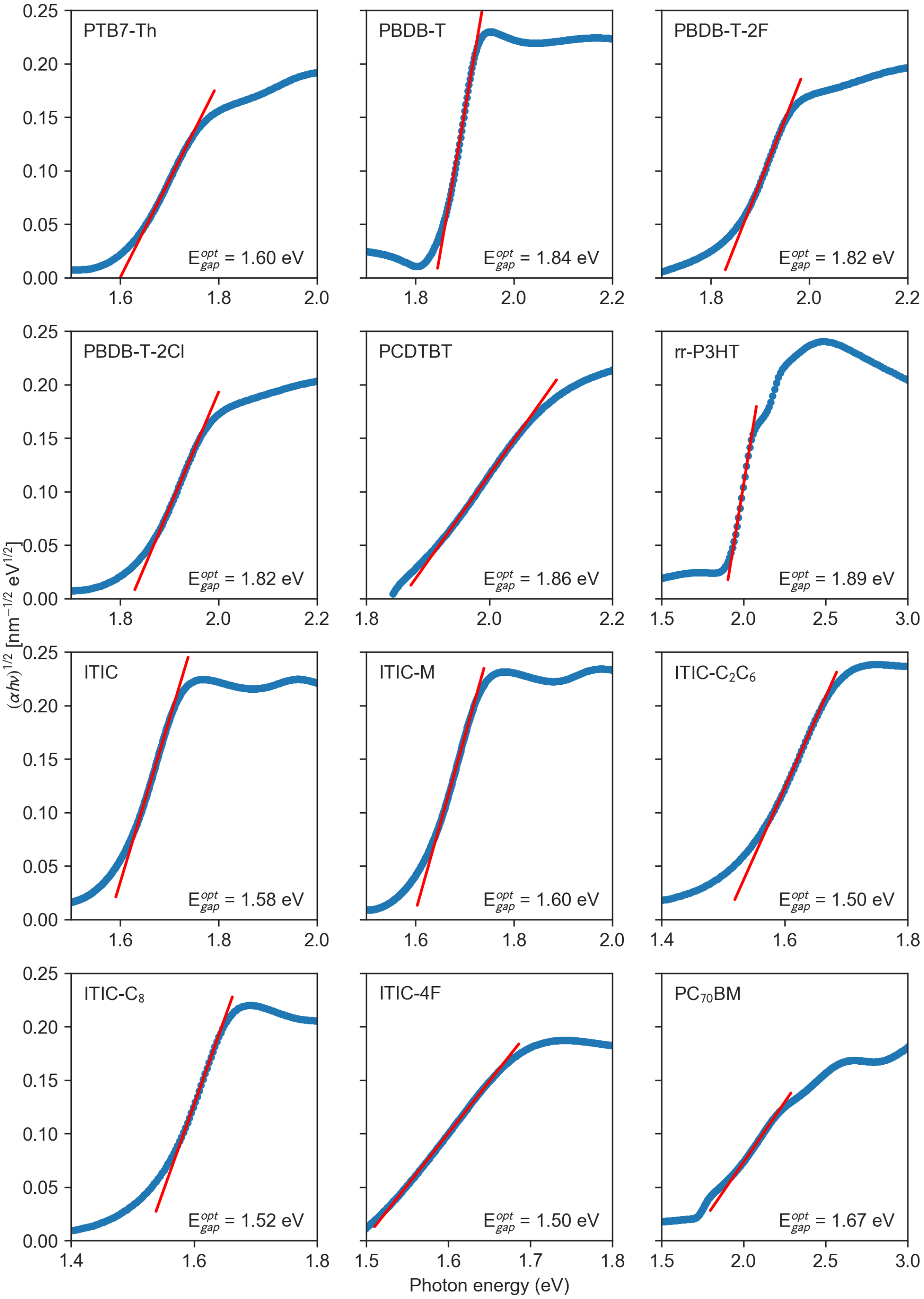


Figure S13. Tauc plots used to determine the solid-state optical band gap of the semiconducting materials studied.

Section IV. Listing of the optoelectronic descriptors used in the AI modelling

Donor and acceptor materials are characterized by a set of fundamental and intrinsic optoelectronic descriptors formed by 7 different magnitudes. While the optical descriptors (κ_{max}, λ_{max} and λ_{edge}) are deduced from their corresponding complex refractive indices determined by variable-angle spectroscopic ellipsometry (VASE) (Figure S12), the electronic descriptors (HOMO, LUMO, E_{gap} and $\mu_{h,e}$) are extracted from a variety of literature references. Note that E_{gap} is simply computed as the LUMO-HOMO difference, thus corresponding to the electronic band gap. Therefore, each D:A pair has 14 fundamental optoelectronic descriptors ascribed (7 per material). Below, we detail the values attained in each of the three distinct sets of descriptors selected for the training of AI algorithms.

In the first set (Table S2), we randomly select HOMO and LUMO energy levels from our literature database. We also randomize the selection of the corresponding mobilities. Therefore, we do not carefully select the data to be consistently ascribed to a unique characterization technique (i.e. cyclic voltammetry (CV) or ultraviolet photoelectron spectroscopy (UPS); or space-charge limited current (SCLC) devices or time-of-flight (TOF) measurements).

Table S2. Fundamental optoelectronic features used as a random selection of (electronic) descriptors. κ_{max}, λ_{max} and λ_{edge} were deduced from the ellipsometric measurements shown above.

	Material	κ_{max}	λ_{max} (nm)	λ_{edge} (nm)	HOMO (eV)	LUMO (eV)	E_{gap} (eV)	$\mu_{h,e}$ ($cm^2\ V^{-1}\ s^{-1}$)	Ref.
Donors	PTB7-Th	0.91	625	720	-5.24	-3.66	1.58	1.80E-04	9,10
	PBDB-T	1.37	637	651	-5.33	-3.53	1.80	1.08E-04	11,12
	PBDB-T-2F	0.79	568	642	-5.47	-3.65	1.82	2.97E-04	13
	PBDB-T-2Cl	0.84	567	639	-5.51	-3.57	1.94	2.13E-04	13
	PCDTBT	0.95	549	610	-5.50	-3.60	1.90	9.00E-06	14,15
Acceptors	ITIC	1.60	706	734	-5.64	-3.92	1.72	1.10E-04	1
	ITIC-M	1.68	702	729	-5.58	-3.98	1.60	1.10E-04	11
	ITIC-C_2C_6	1.86	718	756	-5.66	-3.91	1.75	*6.90E-04*	*
	ITIC-C_8	1.68	737	763	-5.63	-3.91	1.72	6.90E-04	1
	ITIC-4F	1.13	724	756	-5.69	-4.07	1.62	5.05E-04	16,17
	$PC_{70}BM$	0.40	479	585	-6.10	-3.90	2.20	1.00E-03	18,19

* For ITIC-C_2C_6, the corresponding HOMO and LUMO values are reported in Section A. The mobility is considered to be the same as ITIC-C_8.

In the second set (Table S3), we hand-pick those articles in which both reported HOMO and LUMO levels were measured by means of CV whenever possible. To further increase the consistency, we also index the value for the carrier mobility reported in the same work (if possible), typically measured by SCLC measurements on single carrier devices.

Table S3. In this selection of optoelectronic descriptors, both HOMO and LUMO energy levels are extracted from references using CV measurements whenever possible. Mobility values correspond preferentially to those reported in the same works as CV measurements. κ_{max}, λ_{max} and λ_{edge} were deduced from the ellipsometric measurements shown above.

	Material	κ_{max}	λ_{max} (nm)	λ_{edge} (nm)	HOMO (eV)	LUMO (eV)	E_{gap} (eV)	$\mu_{h,e}$ ($cm^2\ V^{-1}\ s^{-1}$)	Ref.
Donors	PTB7-Th	0.91	625	720	-5.30	-3.17	2.13	2.40E-04	[20,21]
	PBDB-T	1.37	637	651	-5.33	-3.29	2.04	3.00E-04	[1]
	PBDB-T-2F	0.79	568	642	-5.54	-3.61	1.93	5.00E-04	[22]
	PBDB-T-2Cl	0.84	567	639	-5.52	-3.55	1.97	5.20E-04	[22]
	PCDTBT	0.95	549	610	-5.50	-3.60	1.90	6.20E-05	[14,23]
Acceptors	ITIC	1.60	706	734	-5.64	-3.92	1.72	1.10E-04	[1]
	ITIC-M	1.68	702	729	-5.58	-3.98	1.60	1.10E-04	[11]
	ITIC-C_2C_6	1.86	718	756	-5.66	-3.91	1.75	*6.90E-04*	*
	ITIC-C_8	1.68	737	763	-5.63	-3.91	1.72	6.90E-04	[1]
	ITIC-4F	1.13	724	756	-5.71	-4.15	1.56	4.50E-04	[22]
	PC$_{70}$BM	0.40	479	585	-5.96	-3.90	2.06	5.70E-04	[22,24]

* For ITIC-C_2C_6, the corresponding HOMO and LUMO values are reported in Section A. The mobility is considered to be the same as ITIC-C_8.

In the third set (Table S4), and according to the large dispersion of HOMO/LUMO energy levels as well as mobility values found after accessing more than 80 distinct literature references (Figure S14), we take the median values of each of the electronic descriptors indexed. This is thought to evaluate the sensitivity of the AI algorithms upon usage of third-party measurements.

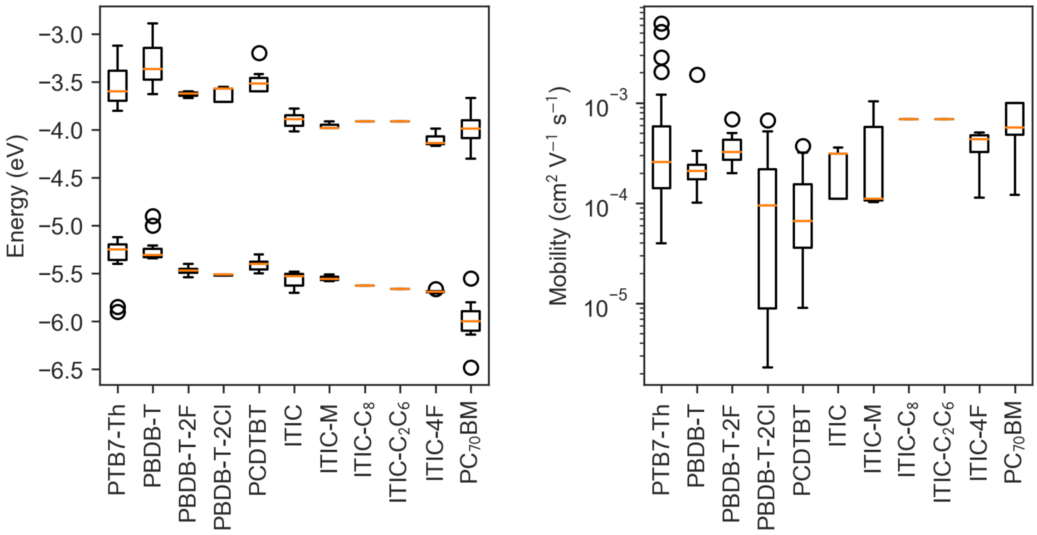


Figure S14. Left panel shows boxplots of the HOMO and LUMO levels for the materials used in this work as collected from the available literature. Right panel illustrates boxplots of the reported mobility extracted by means of SCLC devices in neat and blend films. More than 80 references were indexed in total.

Table S4. Fundamental optoelectronic descriptors obtained as median values after data mining of more than 80 distinct literature references. κ_{max}, λ_{max} and λ_{edge} were deduced from the ellipsometric measurements shown above.

	Material	κ_{max}	λ_{max} (nm)	λ_{edge} (nm)	HOMO (eV)	LUMO (eV)	E_{gap} (eV)	$\mu_{h,e}$ ($cm^2\ V^{-1}\ s^{-1}$)	No. refs.
Donors	PTB7-Th	0.91	625	720	-5.25	-3.60	1.65	2.58E-04	22
	PBDB-T	1.37	637	651	-5.30	-3.37	1.93	2.10E-04	15
	PBDB-T-2F	0.79	568	642	-5.47	-3.62	1.85	3.25E-04	7
	PBDB-T-2Cl	0.84	567	639	-5.51	-3.57	1.94	9.41E-05	5
	PCDTBT	0.95	549	610	-5.40	-3.52	1.88	6.65E-05	10
Acceptors	ITIC	1.60	706	734	-5.53	-3.89	1.64	3.13E-04	8
	ITIC-M	1.68	702	729	-5.56	-3.98	1.58	1.10E-04	3
	ITIC-C_2C_6	1.86	718	756	-5.66	-3.91	1.75	*6.90E-04*	*
	ITIC-C_8	1.68	737	763	-5.63	-3.91	1.72	6.90E-04	1
	ITIC-4F	1.13	724	756	-5.69	-4.14	1.55	4.32E-04	5
	$PC_{70}BM$	0.40	479	585	-6.00	-3.99	2.01	5.70E-04	17

* For ITIC-C_2C_6, the corresponding HOMO and LUMO values are reported in Section A. The mobility is considered to be the same as ITIC-C_8.

From the above lists of fundamental descriptors, an additional series of features is correspondingly derived as detailed in Table S5. These are built from the fundamental magnitudes and might refer to either the neat materials or to their corresponding blend. This second list of derived descriptors contains 9 elements, thus making a total of 23 descriptors per D:A pair (14 + 9). Therein, we identify 8 non-dimensional magnitudes, namely $\kappa_{max,d}$ (*km1*), $\kappa_{max,a}$ (*km2*), $D_{\%S}$ (*dps*), $A_{\%S}$ (*aps*), D%A (*dpa*), CT_e (*ctel*), CT_h (*ctho*) and μ_{imb} (*mamd*).

Table S5. Optoelectronic descriptors derived from any of the fundamental lists of material features provided above.

		Label (*var*)	Definition	Mathematical relationship
Electronic	**Blend**	$E_{gap,d\text{-}a}$ (*eg21*)	Energy gap difference	$E_{gap,d-a} = \lvert E_{gap,d} - E_{gap,a} \rvert$
		Δ_{gap} (*dgap*)	Effective band gap	$\Delta_{gap} = LUMO_a - HOMO_d$
		CT_e (*ctel*)	Charge transfer efficiency for electrons	$CT_e = \frac{\lvert LUMO_d - LUMO_a \rvert}{k_B T}$
		CT_h (*ctho*)	Charge transfer efficiency for holes	$CT_h = \frac{\lvert HOMO_d - HOMO_a \rvert}{k_B T}$
		$\Delta\mu$ (*dmob*)	Difference in mobility	$\Delta\mu = \lvert \mu_d - \mu_a \rvert$
		μ_{imb} (*mamd*)	Mobility ratio	$\mu_{imb} = \frac{\mu_a}{\mu_d}$
Optical	**Neat**	$D_{\%S}$ (*dps*)	Spectral overlap with the sun irradiance	$D_{\%S} = 1 - \frac{2500 - \lambda_{edge,d}}{2500}$
		$A_{\%S}$ (*aps*)	Spectral overlap with the sun irradiance	$A_{\%S} = 1 - \frac{2500 - \lambda_{edge,a}}{2500}$
	Blend	D%A (*dpa*)	Absorption complementarity	$D\%A = 1 - 2\frac{\lambda_{edge,a} - \lambda_{edge,d}}{\lambda_{edge,a} + \lambda_{edge,d}}$

Section V. Training and visualization of the Bayesian machine scientist

In the first run of the Bayesian machine scientist we take the 8 non-dimensional optoelectronic descriptors derived from the fundamental values listed in Table S2 (random selection of features from the literature). For the feeding of the algorithm we randomly select 1,000 data points from the high performing crust of photocurrent values in the corresponding J_{sc}-vol% space of each binary (8 D:A blends in the first run). Note that the upper crust of the J_{sc}-vol% space corresponds to the region where thickness and local morphology are more optimal for the OPV performance. By using a selection of values from the upper crust only, the computational cost is lowered while preserving the meaningful features of the phase space during the training. In addition, the machine scientist is biased to return zero normalized photocurrent (NP) at zero thickness and/or neat films, which gives physical sense to the most plausible model equation eventually found by the algorithm. Such equation returns a mean absolute error (MAE) of 0.09 (±0.02) when trained with the 8-element combinatorial matrix formed by PTB7-Th and PBDB-T blended with ITIC, ITIC-M, ITIC-C_8 and ITIC-C_2C_6 (Figure S15), which includes 8,000 experimental data points in total.

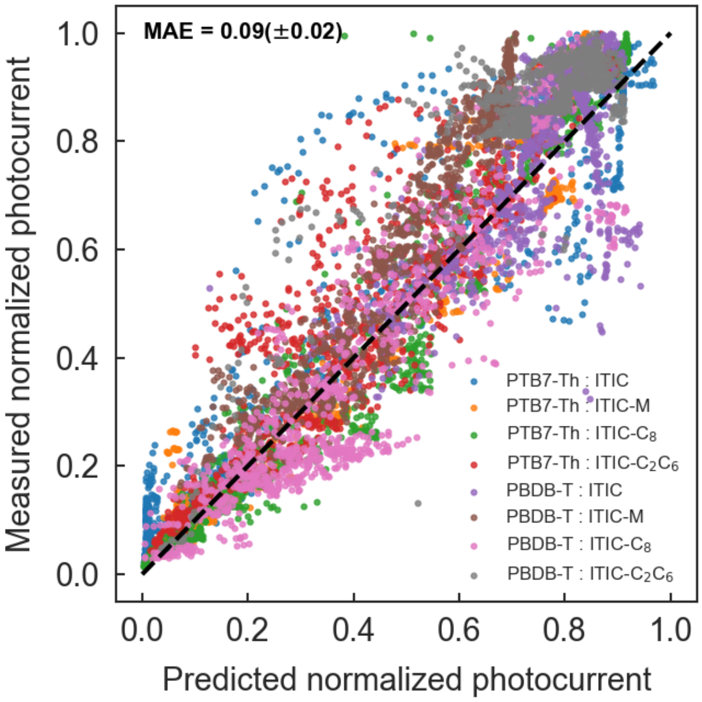


Figure S15. Correlation plot for the normalized photocurrent after training of the Bayesian machine scientist using 8 combinatorial datasets (8,000 experimental data points in total), namely PTB7-Th and PBDB-T individually blended with ITIC, ITIC-M, ITIC-C_8 and ITIC-C_2C_6. The mean absolute error (MAE) found corresponds to 0.09 (±0.02).

The equation found by the Bayesian machine scientist reads:

```
NP = sin((((tan(((((tan(tanh((exp((((((((((((aps / dps) / km2)
** ((_a6_ / ctel) ** (_a3_ / ctho))) / _a9_) * ((DVF **
(tanh(DVF) / (ctho + (DVF + km2)))) ** ((_a8_ + ALT) * dpa))) +
(_a7_ * ALT)) + ((DVF * _a3_) / (_a4_ + dpa))) + aps) + ((_a1_ +
DVF) / dpa)) / _a3_) + (aps * (_a5_ * (ctel + ALT))))) * (km1 +
ctel)))) * ((DVF / mamd) + (((ALT ** dps) * (mamd + ((_a3_ +
(_a7_ ** _a7_)) ** _a0_))) + _a9_))) / km2) * _a2_) / tan(dps)))
+ _a2_) / ((km2 ** km2) ** pow3(DVF))) * dpa)),
```

This equation includes two variables, namely the active layer thickness (ALT, in nm) and the donor volume fraction (DVF); and ten free parameters, valued as:

```
'_a0_' = 841.0242941013105,

'_a1_' = 0.9250910180810115,

'_a2_' = 0.006865810969620029,

'_a3_' = 0.20851084069638415,

'_a4_' = -1.0786072008599603,

'_a5_' = -1.6604993723756856,

'_a6_' = 7.989472231285854e+34,

'_a7_' = 0.10003869806486294,

'_a8_' = 77.87116793766913,

'_a9_' = -0.11596721509342561.
```

From the ensemble of plausible equations found by the machine scientist, the algorithm quantifies the importance of the features according to their rate of appearance. For example: a descriptor showing a feature importance (FI) of 0.50 implies that such parameter appears in 50% of the most plausible model equations found by the algorithm. Since the machine scientist finds quite complicated mathematical relationships, it ends up using most variables most of the times, thus leading to an uninformative classification of variables:

```
'km1' = 0.89, 'km2' = 1.00, 'aps' = 0.96, 'ctel' = 0.93, 'ctho'
= 0.87, 'dpa' = 0.87, 'dps' = 0.93, 'mamd' = 0.87.
```

Nevertheless, and as shown in Figure S16, the model equation demonstrates capability to complement the experimental high-throughput exploration and fill the corresponding photocurrent phase space with mild patterns.

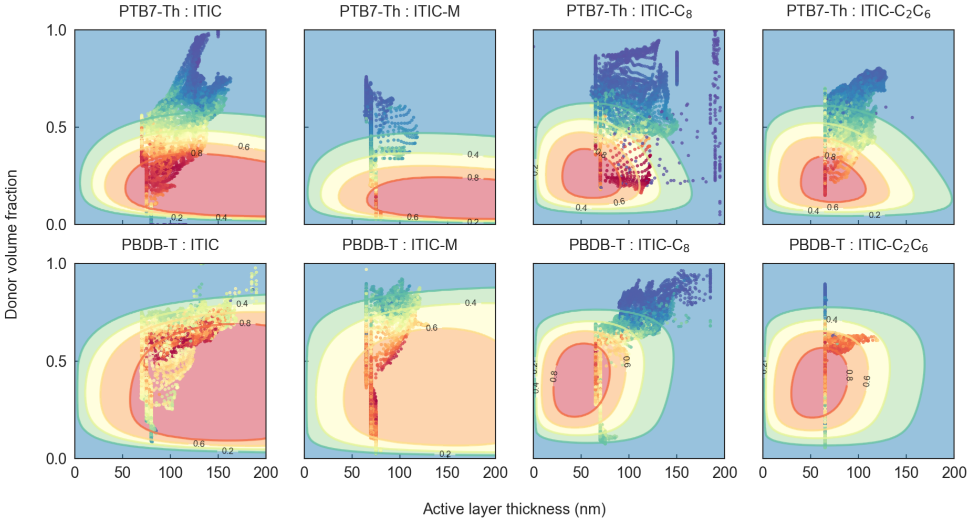


Figure S16. Contour plots for the normalized photocurrent model equation found by the Bayesian machine scientist. In this case, the scattered data includes all the experimentally accessed data points obtained in each combinatorial D:A device.

Contrary to its remarkable descriptive power, the model equation does not extrapolate well in unknown D:A pairs. In this regard, we illustrate in Figure S17 the predicted photocurrent phase space in four different binary validation datasets, namely PTB7-Th:ITIC-4F, PBDB-T-2Cl:ITIC-4F, PBDB-T:PC$_{70}$BM and PCDTBT:PC$_{70}$BM. The equation clearly mismatches the experimental normalized photocurrent both in absolute values and overall trends. Therefore, the machine scientist cannot be employed as an accurate predictive model out of the training set, at least when the training set is limited to the aforementioned 8 distinct D:A blends.

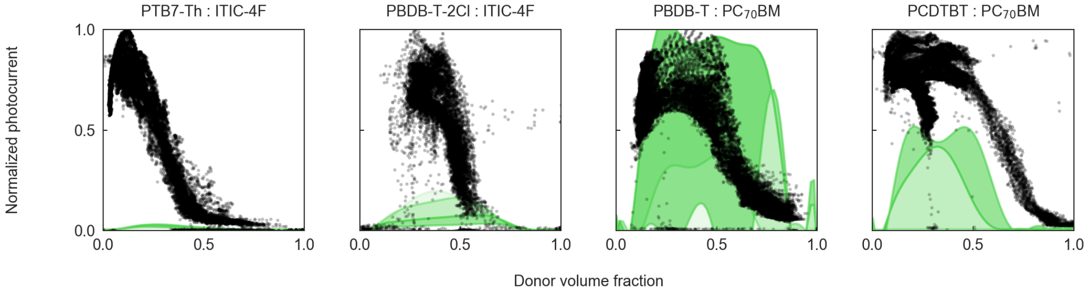


Figure S17. Model predictions drawn by the Bayesian machine scientist in four D:A pairs out of the training set. The experimental scattered data (black dots) are not properly reproduced by the robot (green curves and shaded areas) within the explored phase space diagram, i.e. from neat films up to 200 nm in active layer thickness.

We then employed the median values of the E$_{gap}$s to train a second Bayesian machine scientist algorithm including this time all the experimentally accessed D:A pairs (15 combinations). For this task, and in order to reduce the computational cost, we selected 500 data points from the high-performing crust of photocurrent values in each of the binary blends explored, thus making up a total of 7500 data points for the training step. After several weeks of training, the resulting 8-parameter, most plausible model equation reads:

```
NP = (-(-(tan((sin((((tanh((((cosh(((sinh(((DVF / _a1_) **
_a1_)) ** (Eg1 + DVF)) * ((-(sqrt(Eg1)) + (Eg1 / pow2(_a6_))) *
((_a1_ * _a3_) * (_a3_ * (_a4_ ** sin(fac(Eg1)))))))) ** (((Eg1
** _a7_) * ((_a4_ + _a4_) + (((_a0_ + ALT) / _a7_) + ((Eg1 +
ALT) + ((Eg1 * (_a7_ + Eg1)) * (((Eg2 * Eg2) * Eg2) +
(cos(((_a2_ + cos((Eg1 ** (Eg2 + Eg2)))) / (_a7_ / fac(Eg2)))) /
cos(Eg1)))))))) ** (_a7_ * (_a3_ * DVF)))) + sin(_a0_)) * ((ALT
* (_a6_ / _a7_)) / (sin(((((_a5_ * Eg2) + _a2_) ** Eg2) + ((Eg2
+ (Eg1 ** (_a1_ ** (sqrt((DVF / (((ALT + ALT) / _a0_) + (_a3_ /
Eg1)))) / (Eg2 ** DVF))))) * _a4_))) + (_a5_ + (pow2((Eg2 /
Eg1)) * Eg2)))))) + (_a5_ * (ALT / log((_a4_ * exp(_a1_)))))) *
Eg1) / (sqrt((tan(sinh(DVF)) ** exp(tan(sinh(DVF))))) + ((ALT +
((-(ALT) + _a2_) / Eg1)) * (_a2_ ** _a6_))))) / (_a1_ **
fac(Eg2)))))) / (((DVF ** _a4_) ** (_a5_ * fac(Eg1))) * Eg1))
```

where

```
'_a0_': 620.2945177575979,

'_a1_': '0.958014069387981,

'_a2_': 3443.5970433103516,

'_a3_': 1.2935329423986563,

'_a4_': 107.62433512234023,

'_a5_': -0.000594846647808181,

'_a6_': -0.971457385756211,

'_a7_': -4.2876008724962595
```

Interestingly, the MAE obtained in this run (Figure S18) is comparable to our previous exploration using a larger set of input descriptors, thus suggesting that the E_{gap}s alone are highly descriptive features for the J_{sc}-vol% dependence.

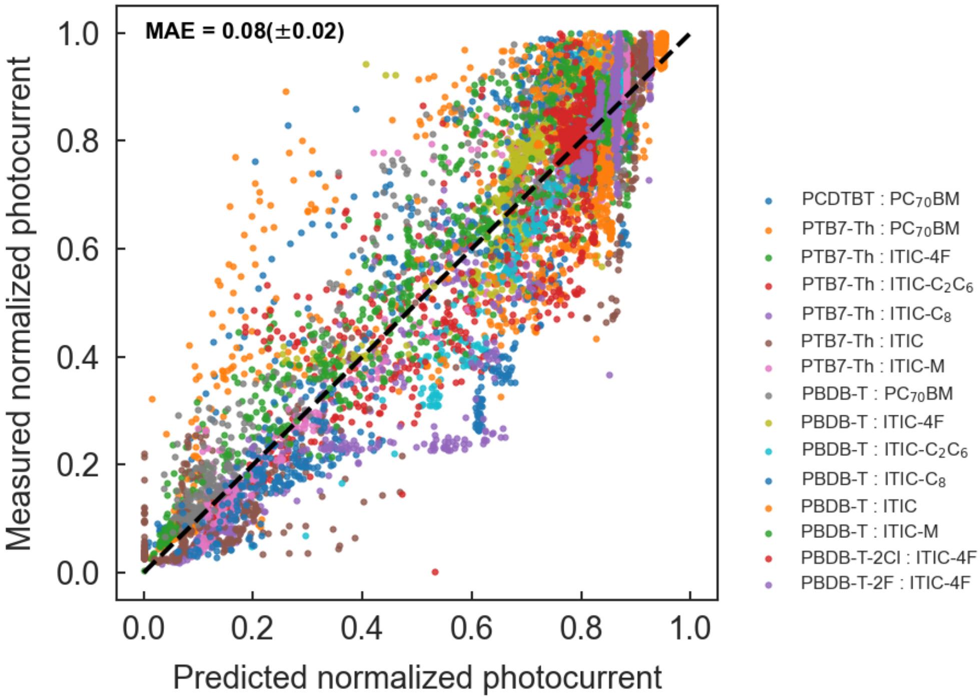


Figure S18. Correlation plot for the normalized photocurrent (NP) obtained after training the Bayesian machine scientist using 15 combinatorial datasets (7500 experimental data points in total), and two material descriptors only, namely the electronic band gaps of donor (*eg1*) and acceptor (*eg2*). The mean absolute error (MAE) found corresponds to 0.08 (±0.02).

By evaluating the abovementioned model equation in a grid, we build largely descriptive contour plots that serve us to follow the mild oscillations of the photocurrent in the thickness-composition space (Figure S19). Finally, we extrapolate the model to four different and unseen D:A pairs to evaluate the predictive capability of the equation found (Figure S20). We observe that the model fails at matching the optimum composition in blends including semicrystalline polymers such as P3HT, similarly to what we observe in the RF model extrapolations (Figure S27).

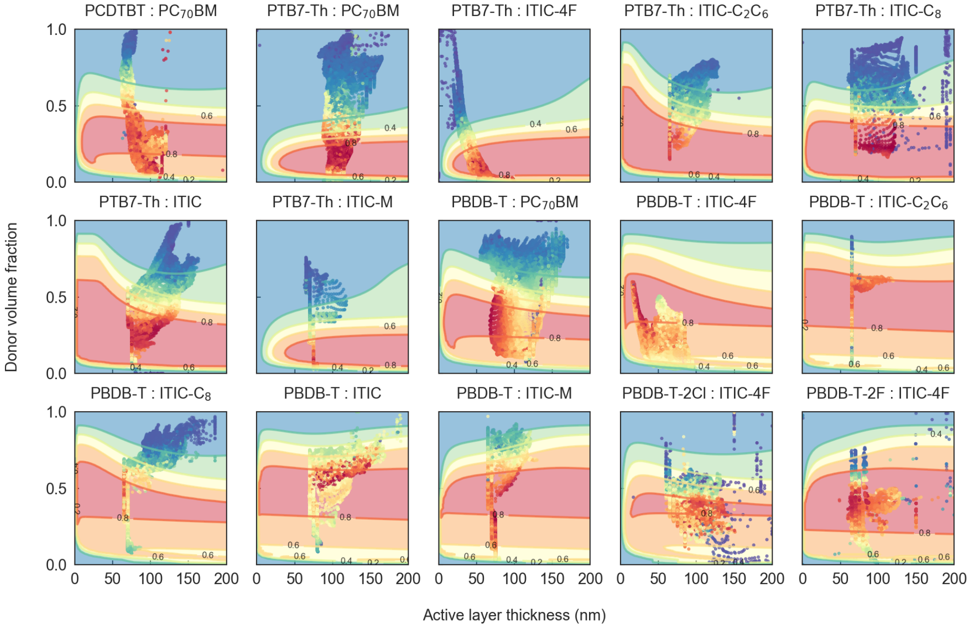


Figure S19. Contour plots for the normalized photocurrent obtained upon training the Bayesian machine scientist using data for 15 distinct D:A combinations (500 data points each) and two input descriptors only, namely the corresponding electronic band gaps of the donor and the acceptor materials. The scattered data shown corresponds to all the data points experimentally generated in each D:A combination, which largely exceed the 500 data points selected for the training step.

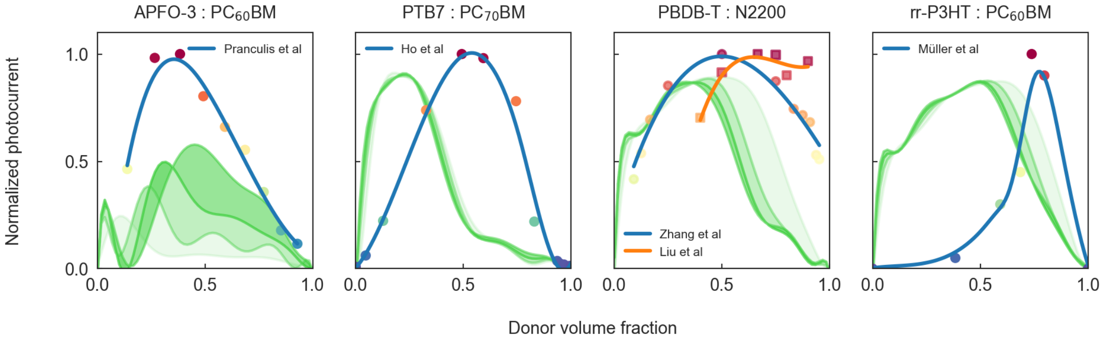


Figure S20. Extrapolations of the model equation found by the Bayesian machine scientist in unseen D:A pairs. Note that the only input parameters are the electronic band gaps of donor and acceptor materials. The active layer thickness dependence on the photocurrent distribution is illustrated by the darkness of the shaded green areas (the darker, the thicker, up to 200 nm).

Section VI. Training of the random forest machine-learning algorithm

a. Random forest ensemble optimization

The RF machine-learning (ML) model is implemented in Python 3.7.3 using the open-source toolbox Scikit-Learn[25] v0.22.2. The RF regressor is first optimized in terms of number of trees in the forest (*n_estimators*) employed by looking for a sweet spot between learning performance and computational cost. The learning performance is evaluated by averaging the mean absolute error (MAE) and mean squared error (MSE) in subsequent leave-one-out cross-validations (LOO-cv) of the 15 experimental binary OPV datasets generated in this work. Our results (Figure S21) show that when all fundamental and optoelectronic descriptors (23 elements) are employed in the training of a RF model, a moderate value of *n_estimators* = 100 yields converged MAE with very limited computational cost (few seconds of training time per LOO-cv). Herein, *n_estimators* is set to 100 by default in our RF training procedures.

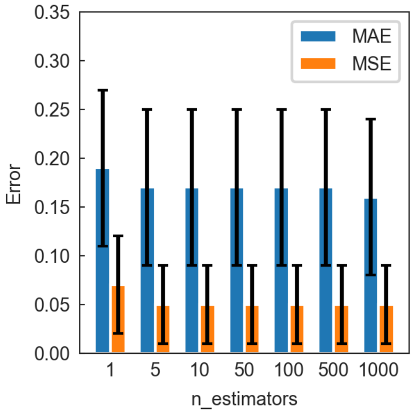


Figure S21. Exploration of the influence of the number of trees in the forest (*n_estimators*) in subsequent LOO-cv runs of 15 combinatorial datasets. The values and error bars depicted account for the mean and standard deviation of the MAE and MSE in the 15 runs performed at each *n_estimators* step.

b. Leave-one-out cross-validations and extrapolation capability

The first tentative LOO-cv included 8 datasets and the full list of optoelectronic descriptors (23). Our (arbitrary) criterion for successful extrapolation required the MAE to be below 0.20 in the validation datasets, which is a figure comparable to our experimental error (10-15%) in the determination of active layer thickness and composition[4]. Based on that threshold error, the RF model yields *ca*. 65% (5/8) accuracy in the present LOO-cv case example (Figure S22). Nevertheless, by extrapolating a single RF model trained with all 8 datasets we obtain a MAE < 0.20 in 7 out of 7 D:A binaries out of the training set (Figure S23). Note that the MAE is computed with respect to the upper shell of 1,000 values in the corresponding J_{sc} vs. active layer thickness and composition phase space (colored rainbow data in Figure

S23). By further extending the training dataset to 15 distinct D:A blends, the LOO-cv returns an average MAE of 0.17 (±0.08), as depicted in Figure S24.

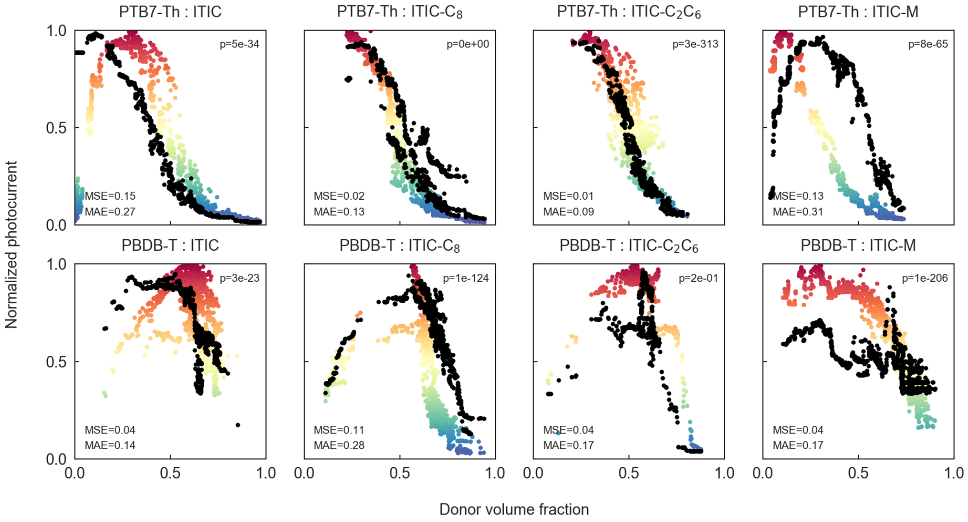


Figure S22. LOO-cv of the RF model trained with 23 descriptors and 8 datasets, namely PTB7-Th and PBDB-T donors blended with ITIC, ITIC-C$_8$, ITIC-C$_2$C$_6$ and ITIC-M. The extrapolation of the RF model is over imposed (black dots) on the actual experimental measurements (in color scale). The corresponding Spearman's rank correlation coefficients (ρ), MAE and MSE obtained in each LOO-cv run are also included.

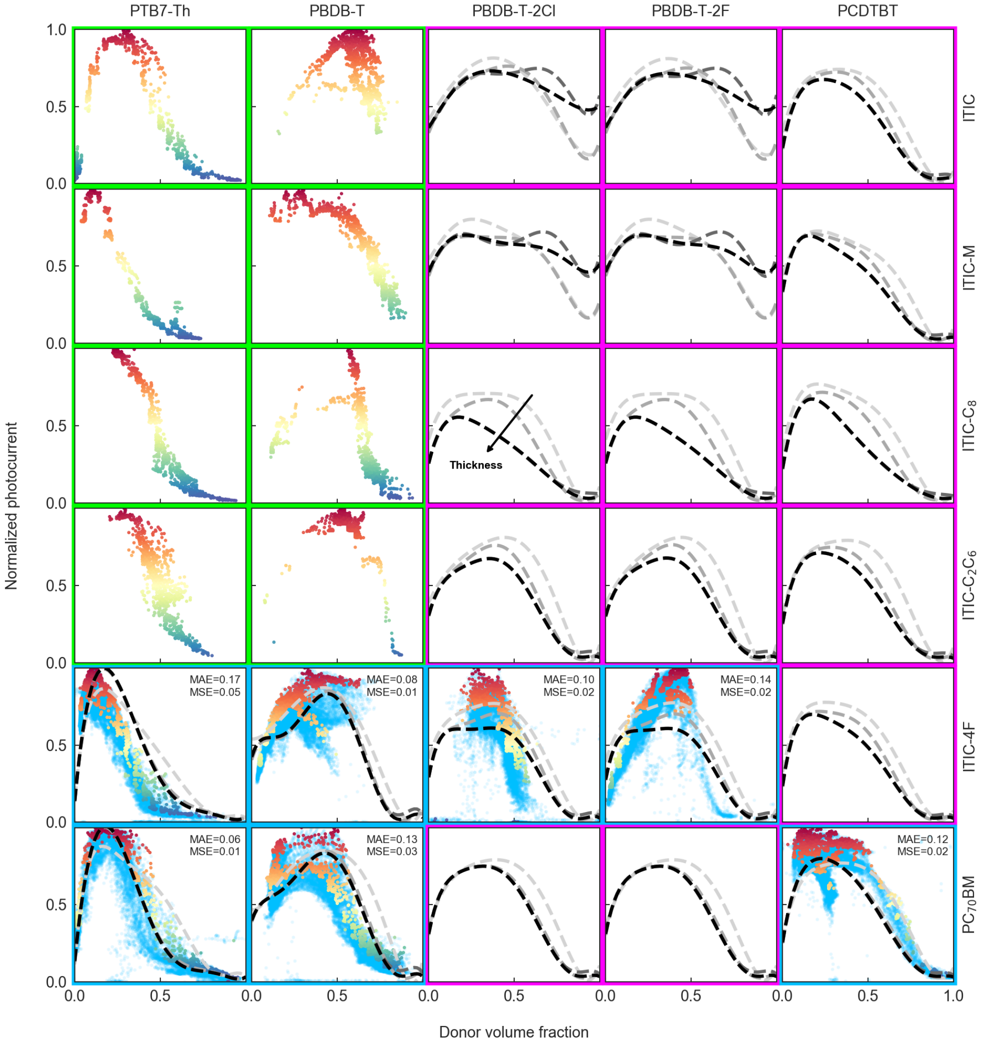


Figure S23. Full combinatorial matrix of the normalized J$_{sc}$ phase diagram of five donor polymers and six different small molecule acceptors built using a RF model trained with 8 datasets (highlighted in green) and 23 optoelectronic descriptors. Subplots highlighted in blue are used to test the predictive capability of the model (they were not used for training in the present case). Magenta panels remark those binaries in which there are not experimental data available, thus they are purely predictive cases. The RF model predictions are depicted as dashed lines corresponding to four different active layer thickness values: 50 nm (lightest grey), 100 nm, 150 nm and 200 nm (black).

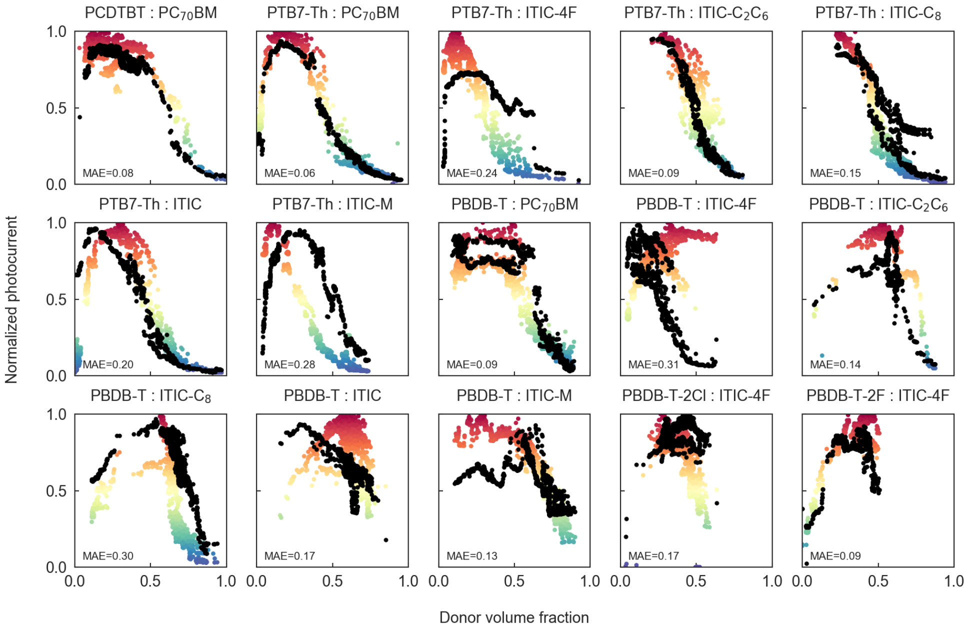


Figure S24. LOO-cv of the RF model trained with 23 descriptors and 15 datasets. The extrapolation of the RF model is over imposed (black dots) on the actual experimental measurements (in color scale). The average MAE obtained is 0.17 (±0.08).

c. Feature selection

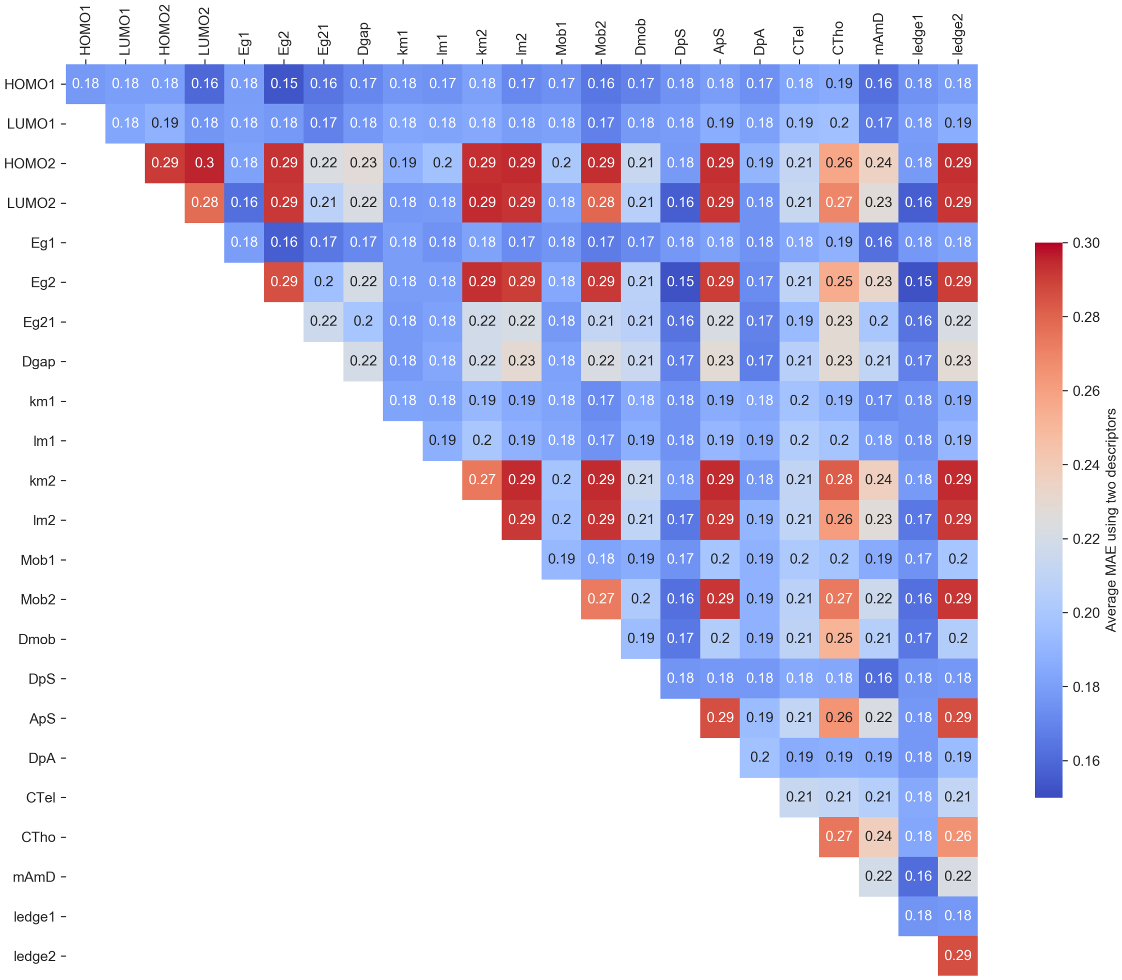
Feature selection in ML algorithms such as RF consists of identifying those descriptors that best describe the given set of observables. This is useful to improve the accuracy of the model while avoiding overfitting and reducing the computational cost. Thus, feature selection generally results in simpler and more generalized models which might enable the understanding of their underlying learning structure to perform an intuitive interpretation, i.e. providing some physical sense to the model predictions in terms of normalized J_{SC} for a variety of material systems. The here proposed initial list of optoelectronic descriptors is notably large (23) yet insignificant with respect to the number of observables used in training: 1,000 data points per D:A pair. Note that all experimentally accessed D:A datasets (15 OPV binaries) are employed in the feature selection procedure to maximize the parametric variability *seen* by the model, thus making up a total of 14,000 data points for the training step (in subsequent LOO-cv runs). This implies that overfitting is not expected to arise; therefore, in our case feature selection is performed to identify the most relevant descriptors and to provide physical intuition to the learning algorithm.

Feature selection in ML does not follow a unique recipe or searching protocol. Conversely, it constitutes a route open to distinct approaches and statistical criteria depending on the type of problem (classification vs regression, categorical vs numerical features) and the information collected *a priori* about the target features in the explored datasets. In this work, we adopted a greedy mean absolute error (G-MAE) feature selection procedure. This approach starts by considering the simplest RF model (i.e. excluding any feature in the modeling) and progressively screens the MAE of models of increased complexity. In our case, the G-MAE exploration was subsequently performed by LOO-cv of 15 datasets using each of the optoelectronic descriptors selections detailed in Section IV. We then averaged the MAE obtained for each combination of descriptors to build a G-MAE matrix (Figure S25). In this way, we also consider the variability of the input descriptors took from the literature. By following this protocol, we identify a series of RF models of two features only that minimize the MAE to *ca*. 0.15, which is a lower figure compared to that obtained using the full list of 23 descriptors (Figure S24).

Among the most successful parametric combinations, we generally observe that descriptors related to the HOMO energy level of the donor such as *dps* or *ledge1* (as well as *homo1* itself) result in accurate models. Regarding acceptor material descriptors, $E_{gap,a}$ (*eg2*) appears as the most successful feature in RF model training. Also, the combination of μ_{imb} (*mamd*) with features related to the HOMO of the donor lead to models with low MAE.

With all that, one of the most interesting parametric combinations identified is that formed by the donor and acceptor materials band gaps, $E_{gap,d}$ and $E_{gap,a}$. The intuitive understanding of such model is remarkable since it indicates that the RF model requires only the electronic band gaps of the mixed components to predict the J_{sc}-vol% phase space. Thus, the complementary absorption between donor and acceptor, determined by their respective band gaps, is likely to be enough for the RF ensemble to shape the normalized J_{sc} phase space of the corresponding blends.

Figure S27 illustrates the RF model extrapolations obtained in binary OPV systems out of the training set formed by the 15 D:A combinations explored in this work, using the electronic band gaps as the only input descriptors. The experimental datasets shown in Figure S27 were extracted from literature[26–30] and include several examples of polymer:fullerene blends as well as a representative example of an all-polymer binary blend, namely PBDB-T:N2200. Notably, in two of the blends explored (APFO-3:PC$_{60}$BM and PBDB-T:N2200) the RF model is close to reproducing the position of the normalized J_{sc} maximum. Nevertheless, the model clearly fails in predicting the J_{sc}-vol% dependence of workhorse D:A pairs such as P3HT:PC$_{60}$BM. This might be indicative of limited extrapolation capabilities when the morphology of the blended materials widely differs with respect to the systems explored in the training step, namely mostly amorphous vs. semi-crystalline polymers as donors. However, further extension of the training dataset to such semi-crystalline systems could enhance the predictive power of the RF model.



	HOMO1	LUMO1	HOMO2	LUMO2	Eg1	Eg2	Eg21	Dgap	km1	lm1	km2	lm2	Mob1	Mob2	Dmob	DpS	ApS	DpA	CTel	CTho	mAmD	ledge1	ledge2
HOMO1	0.18	0.18	0.18	0.16	0.18	0.15	0.16	0.17	0.18	0.17	0.18	0.17	0.17	0.16	0.17	0.18	0.18	0.17	0.18	0.19	0.16	0.18	0.18
LUMO1		0.18	0.19	0.18	0.18	0.18	0.17	0.18	0.18	0.18	0.18	0.18	0.18	0.17	0.18	0.18	0.19	0.18	0.19	0.2	0.17	0.18	0.19
HOMO2			0.29	0.3	0.18	0.29	0.22	0.23	0.19	0.2	0.29	0.29	0.2	0.29	0.21	0.18	0.29	0.19	0.21	0.26	0.24	0.18	0.29
LUMO2				0.28	0.16	0.29	0.21	0.22	0.18	0.18	0.29	0.29	0.18	0.28	0.21	0.16	0.29	0.18	0.21	0.27	0.23	0.16	0.29
Eg1					0.18	0.16	0.17	0.17	0.18	0.18	0.18	0.17	0.18	0.17	0.17	0.18	0.18	0.18	0.18	0.19	0.16	0.18	0.18
Eg2						0.29	0.2	0.22	0.18	0.18	0.29	0.29	0.18	0.29	0.21	0.15	0.29	0.17	0.21	0.25	0.23	0.15	0.29
Eg21							0.22	0.2	0.18	0.18	0.22	0.22	0.18	0.21	0.21	0.16	0.22	0.17	0.19	0.23	0.2	0.16	0.22
Dgap								0.22	0.18	0.18	0.22	0.23	0.18	0.22	0.21	0.17	0.23	0.17	0.21	0.23	0.21	0.17	0.23
km1									0.18	0.18	0.19	0.19	0.18	0.17	0.18	0.18	0.19	0.18	0.2	0.19	0.17	0.18	0.19
lm1										0.19	0.2	0.19	0.18	0.17	0.19	0.18	0.19	0.19	0.2	0.2	0.18	0.18	0.19
km2											0.27	0.29	0.2	0.29	0.21	0.18	0.29	0.18	0.21	0.28	0.24	0.18	0.29
lm2												0.29	0.2	0.29	0.21	0.17	0.29	0.19	0.21	0.26	0.23	0.17	0.29
Mob1													0.19	0.18	0.19	0.17	0.2	0.19	0.2	0.2	0.19	0.17	0.2
Mob2														0.27	0.2	0.16	0.29	0.19	0.21	0.27	0.22	0.16	0.29
Dmob															0.19	0.17	0.2	0.19	0.21	0.25	0.21	0.17	0.2
DpS																0.18	0.18	0.18	0.18	0.18	0.16	0.18	0.18
ApS																	0.29	0.19	0.21	0.26	0.22	0.18	0.29
DpA																		0.2	0.19	0.19	0.19	0.18	0.19
CTel																			0.21	0.21	0.21	0.18	0.21
CTho																				0.27	0.24	0.18	0.26
mAmD																					0.22	0.16	0.22
ledge1																						0.18	0.18
ledge2																							0.29

Figure S25. G-MAE matrix of RF models trained with combinations of two descriptors only using 15 experimental datasets. The MAE values were obtained by averaging the MAE in 15 successive LOO-cv runs for each of the optoelectronic descriptors selections detailed in Section IV. The diagonal components correspond to the MAE obtained in one-parameter RF models. The donor volume fraction (DVF) and the film thickness (FT) are included implicitly in all models.

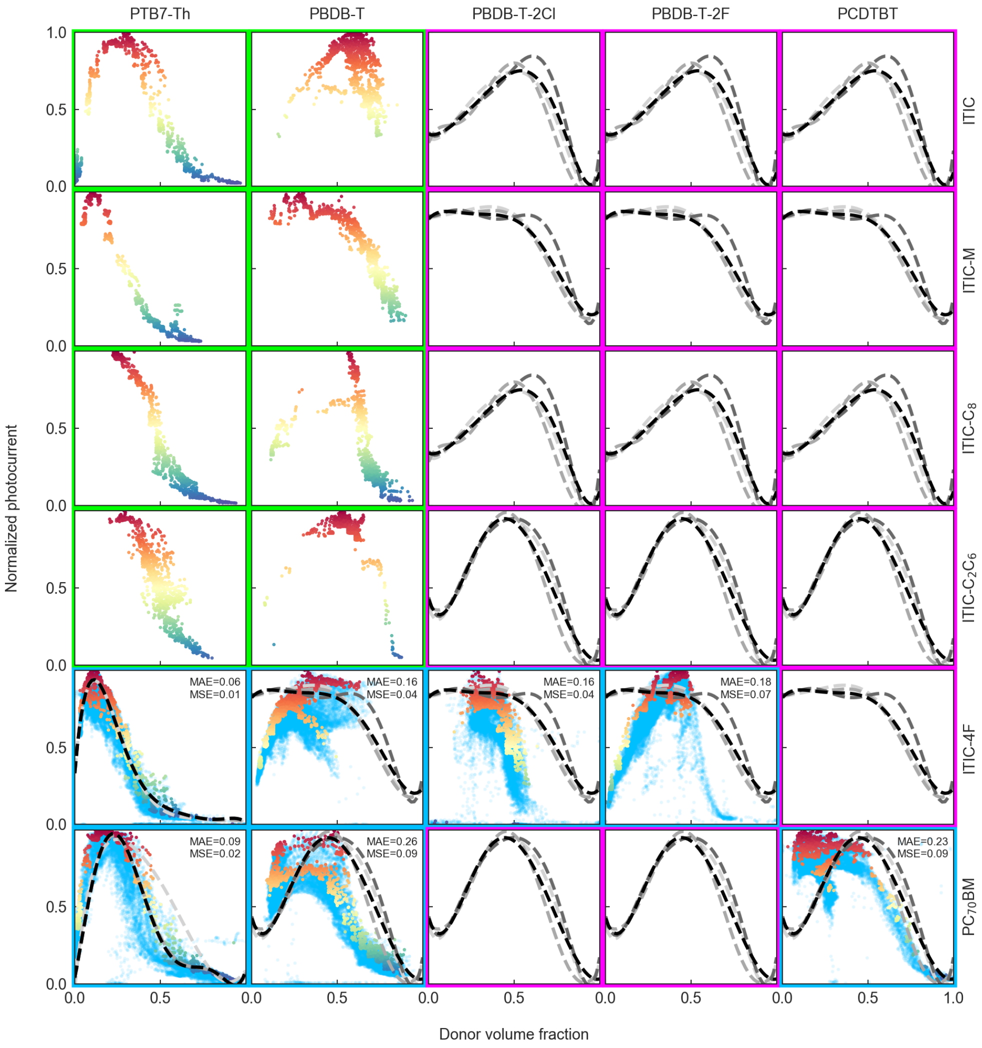


Figure S26. Full combinatorial matrix obtained after RF model training with 8 datasets (green frames) and 2 descriptors only, namely *eg1* and *eg2*. By constraining the extension of the training dataset, the MAE raises up to 0.16 (±0.07).

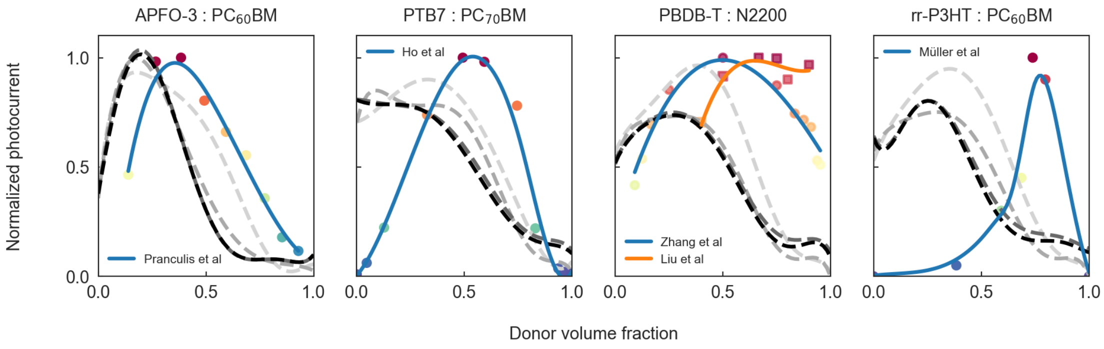


Figure S27. RF model extrapolations in binary OPV blends out of the initial training set. For the training of the RF, only two features were considered, namely *eg1* and *eg2*, apart from DVF and FT. The full list of D:A blends explored in this work (15) was employed in the training step. The scattered experimental datasets depicted were extracted from literature.[26–30] In the conversion from wt% to vol%, an average density for the polymers and fullerenes of 1.1 g mL^{-1} and 1.6 g mL^{-1} was considered, respectively.

d. Sensitivity analysis of the input descriptors in the two-parameter model

In this work, the HOMO, LUMO (thus E_{gap}) and mobility values are not consistently measured in the neat materials; they are, instead, took from a variety of literature references. Since there might be large quantitative differences between characterization methods (Section IV, Supporting Information), it is interesting to evaluate how critical these values are for the successful training of a RF model.

For that purpose, we trained a RF ensemble using the median values obtained in our HOMO/LUMO data mining study (Figure S14). As shown in Figure S28, the median values obtained in the corresponding distributions are still largely descriptive inputs for the RF algorithm. The model returns a MAE of 0.16 (±0.07), a figure which is in excellent agreement with our previous models using hand-picked literature values from CV measurements only (Fig. 6 in the main manuscript).

Finally, since Tauc plots (Figure S13) can be employed to quantify the solid-state optical band gap of thin film materials, we further validated a RF model using such descriptors for the donor and acceptor materials only ($E_{g,opt,d}$ and $E_{g,opt,a}$). This approach is thought to keep the experimental consistency in our work since the complex refractive indices of the here studied materials were all measured in our group using the same experimental setup (a Semilab GES5E rotating polarizer ellipsometer). With all that, the results shown in Figure S29 indicate once again great modeling consistency as the MAE slightly raised to 0.18 (±0.07).

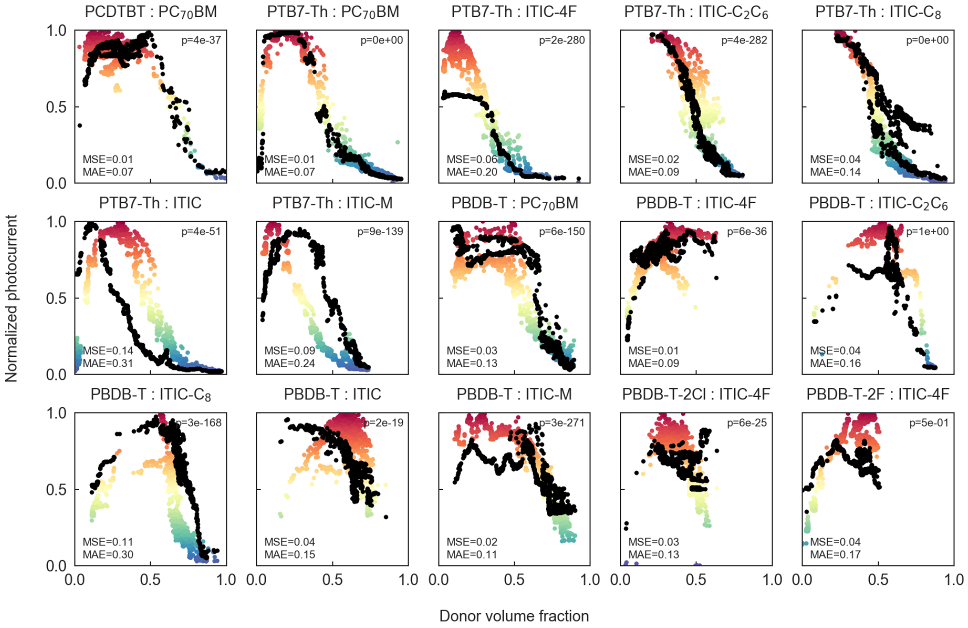


Figure S28. LOO-cv of a 2-parameter RF model trained using the median values of the electronic band gaps found in the literature. The resulting MAE reads 0.16 (±0.07).

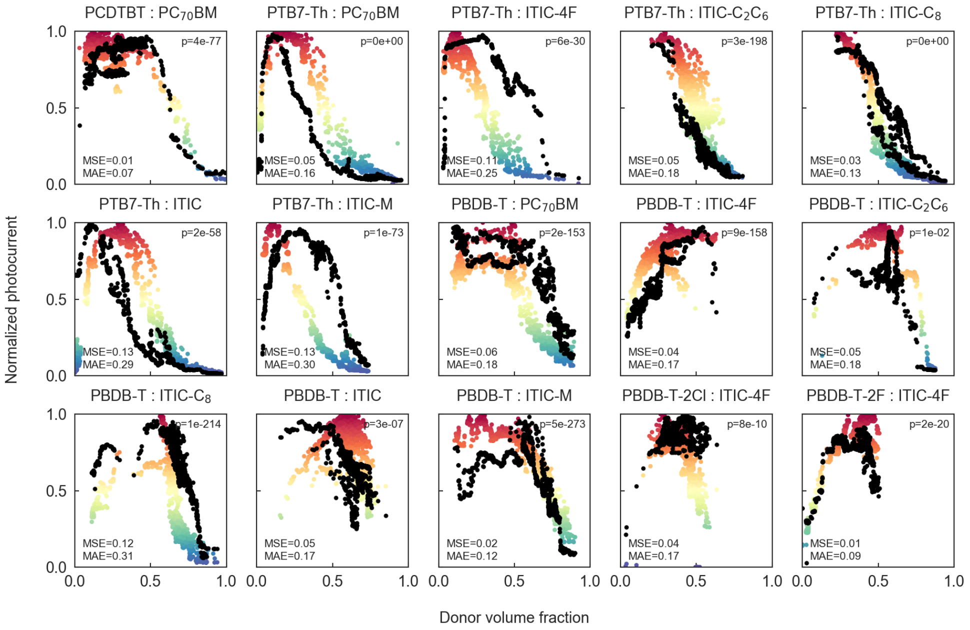


Figure S29. LOO-cv of a 2-parameter RF model trained using as input descriptors the corresponding solid-state optical band gaps of the materials, as obtained from the Tauc plots shown in Figure S13. The average MAE obtained equals 0.18 (±0.07).

Section VII. Conversion from volume to weight fractions via density ratio

Optical probes such as Raman spectroscopy, UV-vis photometry or ellipsometry have access to volume fractions and ratios as these are the fundamental properties associated to effective medium approximations of mixtures. However, the conversion from volume fractions (v_d and v_a for donor and acceptor, respectively) to weight fractions (w_d, w_a) in a blend is straightforward if the density ratio of the blended materials ($\xi = \rho_d / \rho_a$) is known by applying the following formula:

$$w_a = \frac{v_a \rho_a}{\rho_d + v_a(\rho_a - \rho_d)} = 1 - w_d$$

By exploiting our series of discrete pixels of controlled D:A ratios used originally to validate the high-throughput methodology, we can actually determine ξ by fitting the equation above to our experimental data. Interestingly, there is no need to determine the absolute values of ρ_d and ρ_a since ξ alone controls the conversion from volume to mass fractions (and vice versa). Figure S30 shows the volume fraction as deduced from Raman in homogeneous samples of controlled D:A ratios (w:w) for PTB7-Th:ITIC and PBDB-T:ITIC-C$_2$C$_6$ blends. These were prepared by pipetting controlled weights of a pristine acceptor aliquot in distinct vials containing neat polymer only; this procedure minimizes the error committed during pipetting viscous inks (neat polymers). A strong deviation from a straight line in Figure S30 indicates the existence of a large density difference (see green and red dashed lines in Figure S30 as extreme case examples).

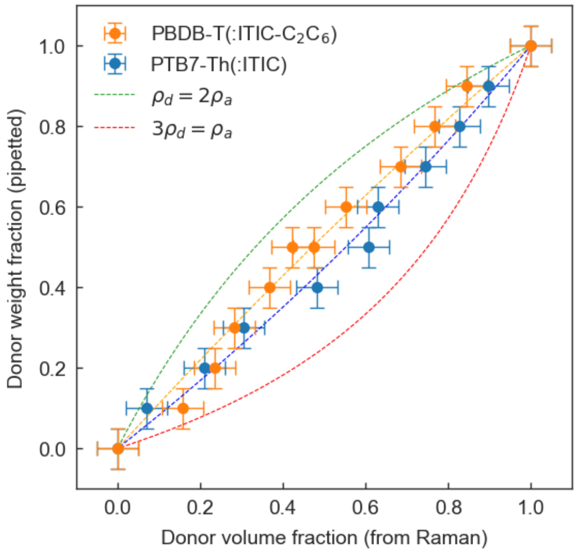


Figure S30. Correlation between donor weight fractions (as prepared from the corresponding aliquots) and donor volume fractions measured by Raman spectroscopy.

In our case, such deviation is not very pronounced, thus suggesting that the D:A material combinations have similar densities. In particular, we have determined ξ = 0.81 and ξ = 1.12 for the PTB7-Th:ITIC and PBDB-T:ITIC-C_2C_6 blends, respectively. Assuming a density for PTB7-Th of 1.17 g cm^{-3} (as reported for PTB7)[31] and a density for PBDB-T of 1.1 g cm^{-3} (similar to other conjugated polymers),[2] the corresponding densities of ITIC and ITIC-C_2C_6 read 1.44 g cm^{-3} and 0.98 g cm^{-3}. With all that, the conversion from volume to weight fractions leads to a lateral displacement of the original volumetric-based data distribution (Figure S31) that is comprised within our experimental error in the determination of composition by Raman spectroscopy (5-10 vol% in the best case).[4]

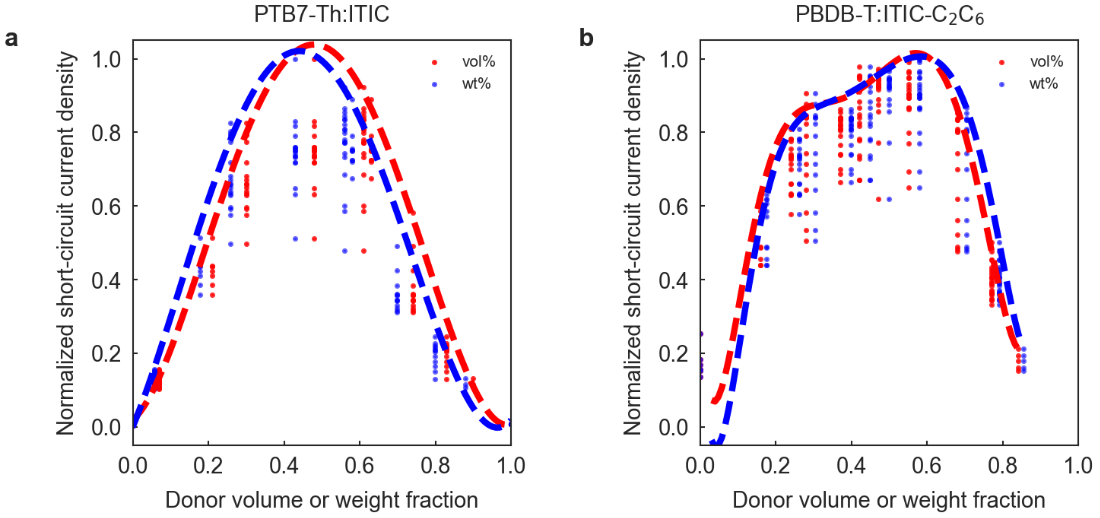


Figure S31. Normalized short-circuit current density distributions obtained in discrete devices of (a) PTB7-Th:ITIC and (b) PBDB-T:ITIC-C_2C_6 blends depicted as a function of donor volume (red) and weight (blue) fractions.

References

1 Z. Fei, F. D. Eisner, X. Jiao, M. Azzouzi, J. A. Röhr, Y. Han, M. Shahid, A. S. R. Chesman, C. D. Easton, C. R. McNeill, T. D. Anthopoulos, J. Nelson and M. Heeney, *Adv. Mater.*, 2018, **30**, 1705209.

2 A. Sánchez-Díaz, X. Rodríguez-Martínez, L. Córcoles-Guija, G. Mora-Martín and M. Campoy-Quiles, *Adv. Electron. Mater.*, 2018, **4**, 1700477.

3 E. Pascual-San-José, X. Rodríguez-Martínez, R. Adel-Abdelaleim, M. Stella, E. Martínez-Ferrero and M. Campoy-Quiles, *J. Mater. Chem. A*, 2019, **7**, 20369–20382.

4 X. Rodríguez-Martínez, M. S. Vezie, X. Shi, I. McCulloch, J. Nelson, A. R. Goñi and M. Campoy-Quiles, *J. Mater. Chem. C*, 2017, **5**, 7270–7282.

5 A. Harillo-Baños, X. Rodríguez-Martínez and M. Campoy-Quiles, *Adv. Energy Mater.*, 2020, **10**, 1902417.

6 S. Hultmark, S. H. K. Paleti, A. Harillo, S. Marina, F. A. A. Nugroho, Y. Liu, L. K. E. Ericsson, R. Li, J. Martín, J. Bergqvist, C. Langhammer, F. Zhang, L. Yu, M. Campoy-Quiles, E. Moons, D. Baran and C. Müller, *Adv. Funct. Mater.*, 2020, **30**, 2005462.

7 W. C. Tsoi, D. T. James, J. S. Kim, P. G. Nicholson, C. E. Murphy, D. D. C. Bradley, J. Nelson and J.-S. Kim, *J. Am. Chem. Soc.*, 2011, **133**, 9834–9843.

8 B. Qi and J. Wang, *Phys. Chem. Chem. Phys.*, 2013, **15**, 8972.

9 S. Zhang, L. Ye, W. Zhao, D. Liu, H. Yao and J. Hou, *Macromolecules*, 2014, **47**, 4653–4659.

10 Y. Tang, H. Sun, Z. Wu, Y. Zhang, G. Zhang, M. Su, X. Zhou, X. Wu, W. Sun, X. Zhang, B. Liu, W. Chen, Q. Liao, H. Y. Woo and X. Guo, *Adv. Sci.*, 2019, **6**, 1901773.

11 S. Li, L. Ye, W. Zhao, S. Zhang, S. Mukherjee, H. Ade and J. Hou, *Adv. Mater.*, 2016, **28**, 9423–9429.

12 W. Zhao, S. Li, S. Zhang, X. Liu and J. Hou, *Adv. Mater.*, 2017, **29**, 1604059.

13 S. Zhang, Y. Qin, J. Zhu and J. Hou, *Adv. Mater.*, 2018, **30**, 1800868.

14 N. Blouin, A. Michaud and M. Leclerc, *Adv. Mater.*, 2007, **19**, 2295–2300.

15 K. K. H. Chan, S. W. Tsang, H. K. H. Lee, F. So and S. K. So, *Org. Electron.*, 2012, **13**, 850–855.

16 Q. Fan, Q. Zhu, Z. Xu, W. Su, J. Chen, J. Wu, X. Guo, W. Ma, M. Zhang and Y. Li, *Nano Energy*, 2018, **48**, 413–420.

17 W. Zhao, S. Li, H. Yao, S. Zhang, Y. Zhang, B. Yang and J. Hou, *J. Am. Chem. Soc.*, 2017, **139**, 7148–7151.

18 Z. Shi, H. Liu, J. Li, F. Wang, Y. Bai, X. Bian, B. Zhang, A. Alsaedi, T. Hayat and Z. Tan, *Sol. Energy Mater. Sol. Cells*, 2018, **180**, 1–9.

19 C. Bracher, H. Yi, N. W. Scarratt, R. Masters, A. J. Pearson, C. Rodenburg, A. Iraqi and D. G. Lidzey, *Org. Electron.*, 2015, **27**, 266–273.

20 C. Cui, W.-Y. Wong and Y. Li, *Energy Environ. Sci.*, 2014, **7**, 2276–2284.

21 Z. Li, X. Xu, W. Zhang, Z. Genene, W. Mammo, A. Yartsev, M. R. Andersson, R. A. J. Janssen and E. Wang, *J. Mater. Chem. A*, 2017, **5**, 11693–11700.

22 M.-A. Pan, T.-K. Lau, Y. Tang, Y.-C. Wu, T. Liu, K. Li, M.-C. Chen, X. Lu, W. Ma and C. Zhan, *J. Mater. Chem. A*, 2019, **7**, 20713–20722.

23 Z. Ding, X. Long, B. Meng, K. Bai, C. Dou, J. Liu and L. Wang, *Nano Energy*, 2017, **32**, 216–224.

24 J. Wang, S. Xie, D. Zhang, R. Wang, Z. Zheng, H. Zhou and Y. Zhang, *J. Mater. Chem. A*, 2018, **6**, 19934–19940.

25 F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, A. Müller, J. Nothman, G. Louppe, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and É. Duchesnay, 2012, arXiv:1201.0490.

26 C. Müller, T. A. M. Ferenczi, M. Campoy-Quiles, J. M. Frost, D. D. C. Bradley, P. Smith, N. Stingelin-Stutzmann and J. Nelson, *Adv. Mater.*, 2008, **20**, 3510–3515.

27 V. Pranculis, Y. Infahsaeng, Z. Tang, A. Devižis, D. A. Vithanage, C. S. Ponseca, O. Inganäs, A. P. Yartsev, V. Gulbinas and V. Sundström, *J. Am. Chem. Soc.*, 2014, **136**, 11331–11338.

28 C. H. Y. Ho, S. H. Cheung, H.-W. Li, K. L. Chiu, Y. Cheng, H. Yin, M. H. Chan, F. So, S.-W. Tsang and S. K. So, *Adv. Energy Mater.*, 2017, **7**, 1602360.

29 Y. Zhang, Y. Xu, M. J. Ford, F. Li, J. Sun, X. Ling, Y. Wang, J. Gu, J. Yuan and W. Ma, *Adv. Energy Mater.*, 2018, **8**, 1800029.

30 X. Liu, Y. Zou, H.-Q. Wang, L. Wang, J. Fang and C. Yang, *ACS Appl. Mater. Interfaces*, 2018, **10**, 38302–38309.

31 O. Dyck, S. Hu, S. Das, J. Keum, K. Xiao, B. Khomami and G. Duscher, *Polymers*, 2015, **7**, 2446–2460.