

## Supplementary Information

### Physics-based extraction of material parameters from perovskite experiments via Bayesian optimization

Hualin Zhan<sup>1</sup>, Viqar Ahmad<sup>1</sup>, Azul Mayon<sup>1</sup>, Grace Dansoa Tabi<sup>1</sup>, Anh Dinh Bui<sup>1</sup>,  
Zhuofeng Li<sup>1</sup>, Daniel Walter<sup>1</sup>, Hieu Nguyen<sup>1</sup>, Klaus Weber<sup>1</sup>, Thomas White<sup>1</sup>, and Kylie  
Catchpole<sup>1</sup>

*<sup>1</sup>School of Engineering, Australian National University, ACT 2601, Australia*

#### Contents

1. Supplementary Notes
2. Supplementary Figures
3. Supplementary Tables
4. Supplementary References

## 1. Supplementary Note

*A quick view of the effect of multi-parameter correlation on parameter extraction.*

During multi-parameter extraction, the calculated cost, i.e., the difference between experimental and theoretical TRPL, is a function of these parameters, for example,  $f=f(x_1, x_2, \dots, x_i, \dots, x_n)$ , where  $x_1, x_2, \dots, x_i, \dots, x_n$  are the parameters to be extracted. As discussed in the main text, our Bayesian optimization (BO) method calculates the differences between the best calculated cost  $f$  so far and the predicted costs  $g$  of all untested parameter values to decide the values of  $x_1, x_2, \dots, x_i, \dots, x_n$  used in the next iterative test, which should be large enough. Here, a predicted  $g$  (by the Gaussian process model) is used because the explicit form of  $f(x_1, x_2, \dots, x_i, \dots, x_n)$  is unknown and the calculation of  $f$  for all values of  $x_1, x_2, \dots, x_i, \dots, x_n$  is impossible.

In an ideal case,  $g$  is close to  $f$  and hence, finding the largest difference between  $f$  and  $g$  for a certain parameter  $x_i$  is analogous to the search of the largest slope of  $f$ , i.e.,  $\frac{df}{dx_i}$ . For the sake of discussion, we write  $\frac{df}{dx_i}$  as

$$\frac{df}{dx_i} = \frac{dx_1 \partial f}{dx_i \partial x_1} + \frac{dx_2 \partial f}{dx_i \partial x_2} + \dots + \frac{\partial f}{\partial x_i} + \dots + \frac{dx_n \partial f}{dx_i \partial x_n}.$$

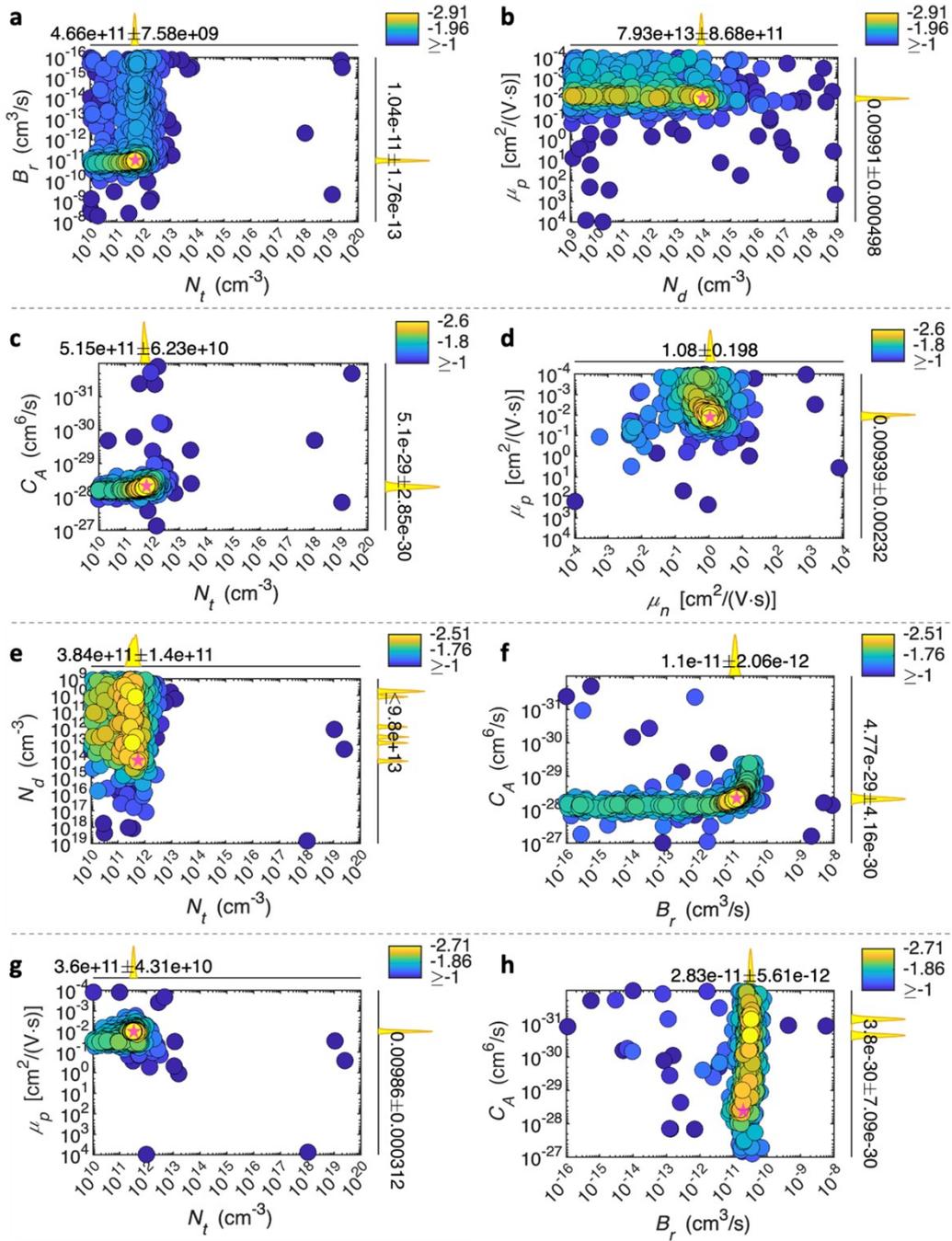
$\frac{dx_1}{dx_i}$  represents the correlation between  $x_1$  and  $x_i$ .  $\frac{\partial f}{\partial x_1}$  is the slope of  $f$  over  $x_1$  when all other parameters are constant. Therefore, the extraction of a parameter  $x_i$  is affected by its correlation with other parameters and the capability to extract other parameters.

Again, we note that BO does not calculate  $\frac{df}{dx_i}$ , partly because the explicit form of  $f$  is unknown and we do not have all values of  $f$  for all possible parameters' values.

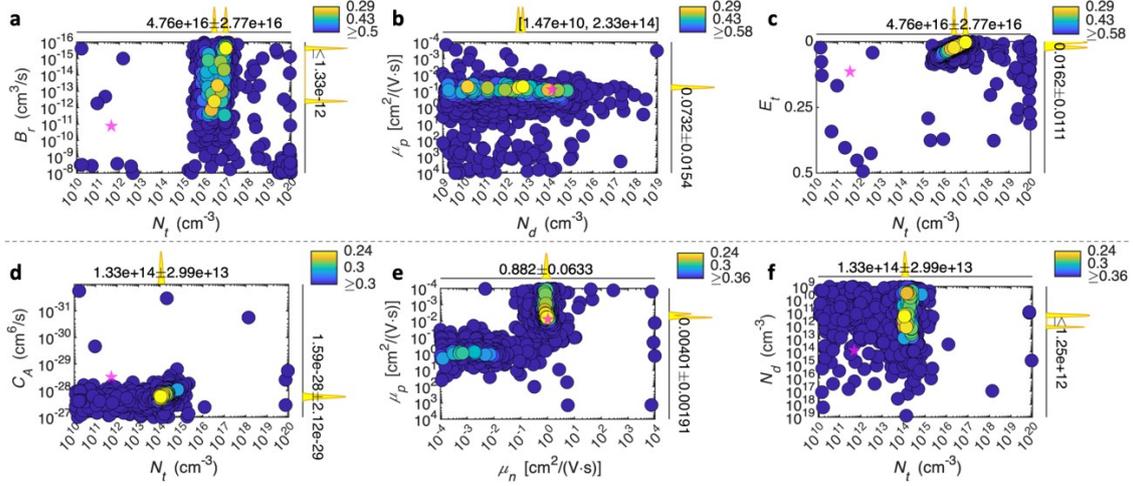
This equation implies that capability of extracting  $x_i$  during a multi-parameter extraction approaches that during a single-parameter extraction when  $\frac{dx_1}{dx_i} \rightarrow 0$ ,  $\frac{dx_2}{dx_i} \rightarrow 0$ , ..., i.e., the

parameter-parameter correlation is minimized, or when  $\frac{\partial f}{\partial x_1} \rightarrow 0$ ,  $\frac{\partial f}{\partial x_2} \rightarrow 0$ , ..., i.e., the local minima form continuous flat lines/planes parallel to all axes except the axis of  $x_i$ .

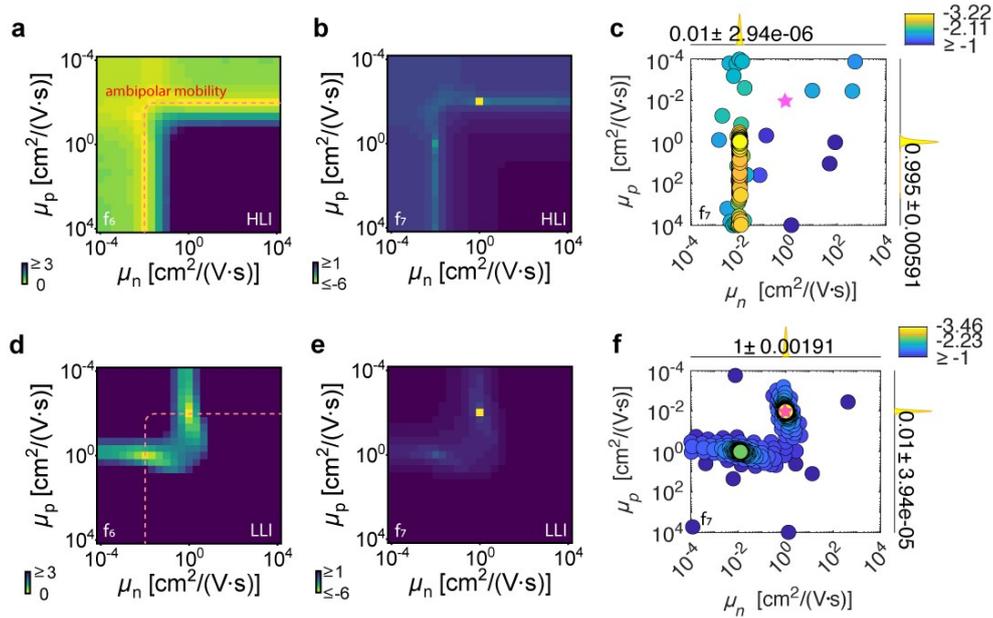
## 2. Supplementary Figures



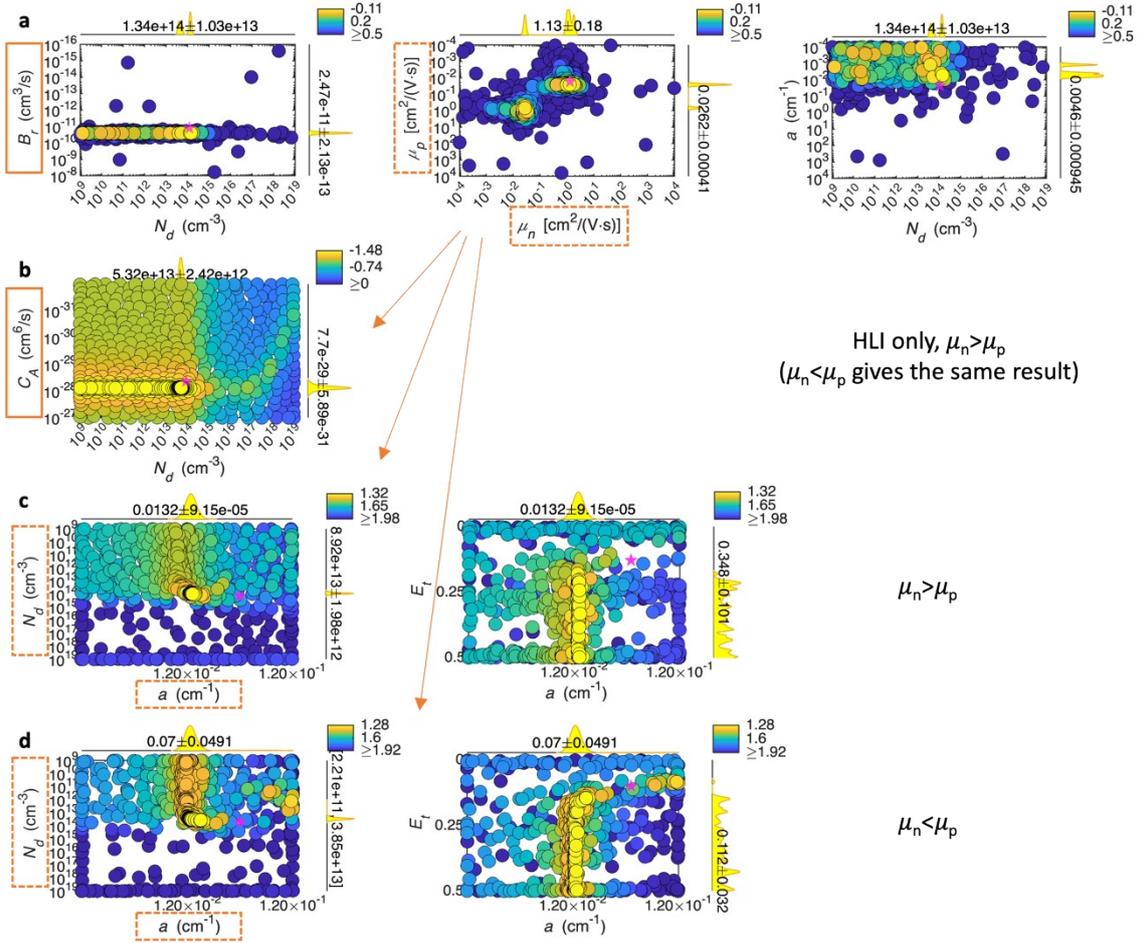
**Fig. S1 | 4D parameter extraction from previously simulated TRPL by direct implementation of BO.** Here four different combinations of four parameters are tested. **a,b**, Successful extraction of  $B_r$ ,  $N_t$ ,  $N_d$ ,  $\mu_p$ . **c,d**, Successful extraction of  $C_A$ ,  $N_t$ ,  $\mu_n$ ,  $\mu_p$ . **e,f**, Imperfect extraction of  $B_r$ ,  $N_t$ ,  $N_d$ ,  $C_A$ , in which the extracted  $N_d$  shows a large uncertainty. **g,h**, Imperfect extraction of  $C_A$ ,  $N_t$ ,  $B_r$ ,  $\mu_p$ , in which the extracted  $C_A$  shows a large uncertainty. The target values are indicated by the magenta stars. Although  $N_d$  and  $C_A$  can be extracted in the combinations of  $(B_r, N_t, N_d, \mu_p)$  and  $(C_A, N_t, \mu_n, \mu_p)$  respectively, they are more difficult to be extracted in the combinations of  $(B_r, N_t, N_d, C_A)$  and  $(C_A, N_t, B_r, \mu_p)$ .



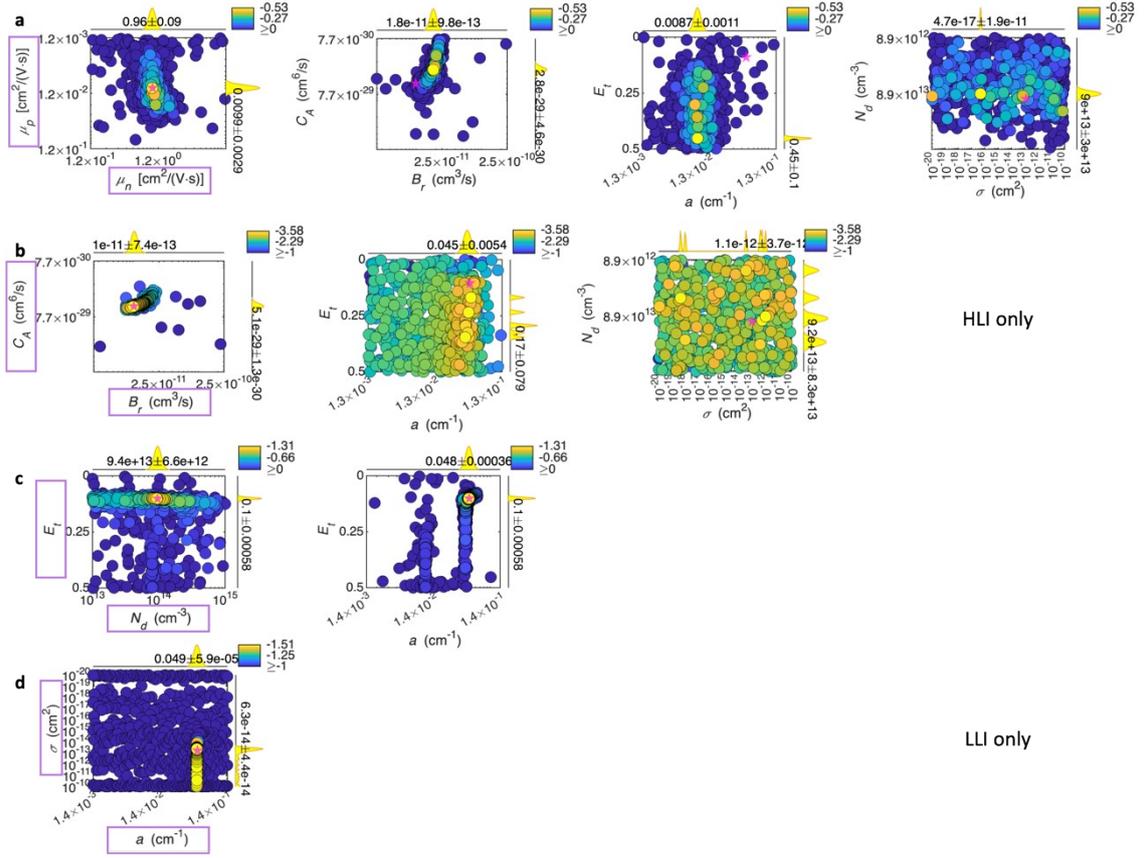
**Fig. S2 | 5D parameter extraction from previously simulated TRPL by direct implementation of BO.** Here we add one extra parameter to the successful 4D parameter extractions in Fig. S1. **a-c**, Attempted extraction of  $B_r$ ,  $N_d$ ,  $\mu_p$ ,  $E_r$ . **d-f**, Attempted extraction of  $C_A$ ,  $N_t$ ,  $\mu_n$ ,  $\mu_p$ ,  $N_d$ . Only the extracted parameters of carrier mobilities are roughly close to target.



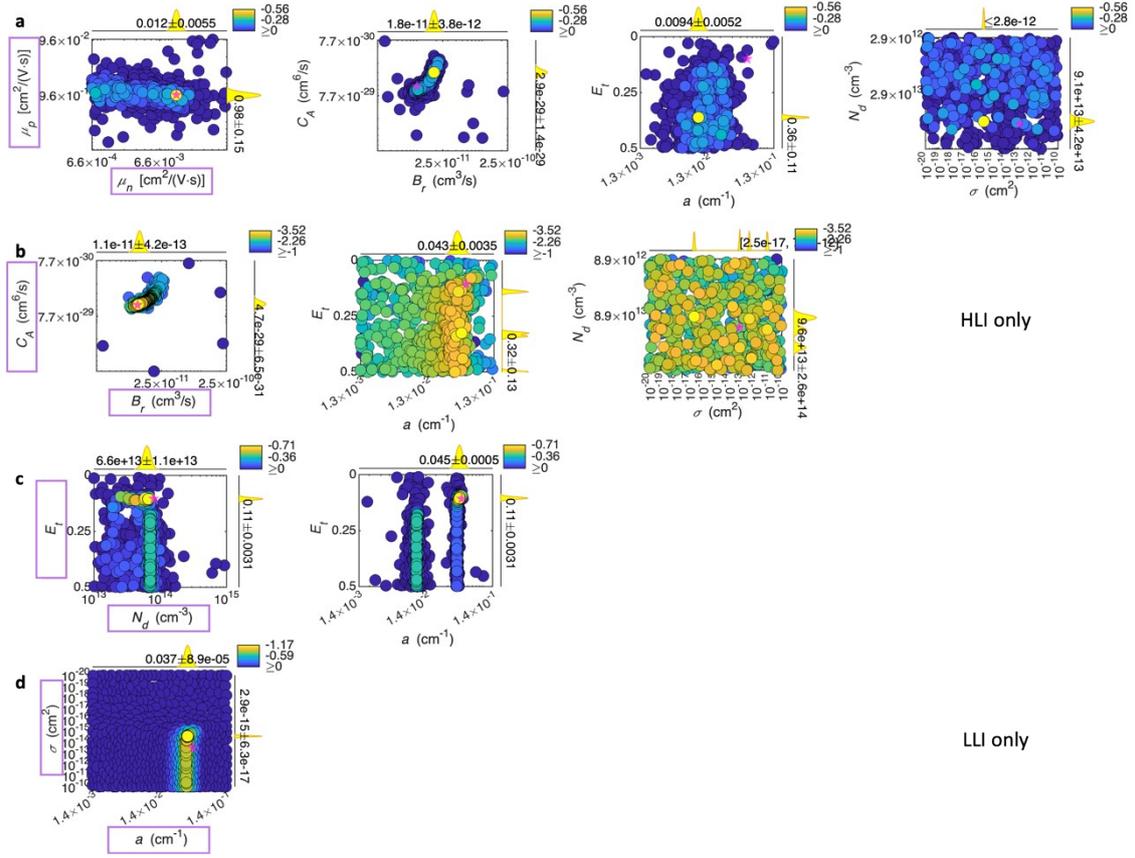
**Fig. S3 | Extraction of carriers' mobilities using different experimental conditions and cost functions.** **a-c**, Grid search (**a,b**) and BO search (**c**) of carriers' mobilities under the HLI condition. An RMSE-type ( $f_6$ ) and a logarithm-type ( $f_7$ ) cost functions are used in (**a**) and (**b**), respectively. The pink dashed lines in (**a,d**) indicate the ambipolar mobility that is derived under the high excess carrier density assumption, i.e., the HLI condition. **d-f**, Grid search (**a,b**) and BO search (**c**) of carriers' mobilities under the LLI condition. Note that, although the BO search is not affected by the ambipolar mobility, BO is possible to extract a set of electron & hole mobility values that is symmetric to the target values. For example, BO may produce  $\mu_n = 0.01$  and  $\mu_p = 1$ , while the actual values are  $\mu_n = 1$  and  $\mu_p = 0.01$ . This uncertainty should be considered in the multi-parameter extraction process by including both symmetric sets of mobility values.



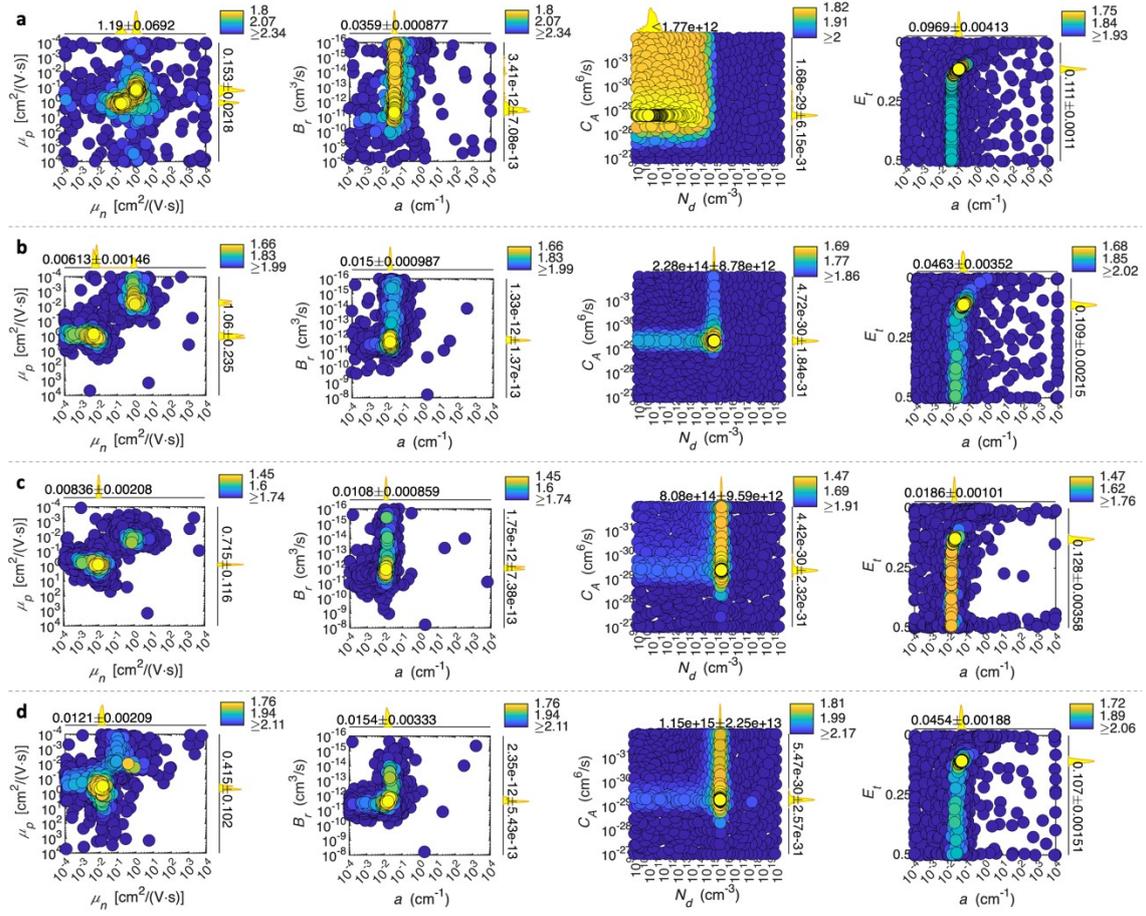
**Fig. S4 | Step 1 of 8D parameter extraction from previously simulated TRPL by our physics-based strategy.** **a**, Estimation of  $B_r$ ,  $\mu_n$ , and  $\mu_p$  by setting  $a$ ,  $B_r$ ,  $N_d$ ,  $\mu_n$ ,  $\mu_p$  as free variables. Two symmetric clusters of points are shown in the middle panel, which are both used in the following steps. **b**, Estimation of  $C_A$  by setting both  $a$  and  $C_A$  as the free variables. **c,d**, Estimation of  $a$  and  $N_d$  by setting  $a$ ,  $N_d$ ,  $E_t$  as the free variables and letting  $\mu_n > \mu_p$  (**c**) and  $\mu_n < \mu_p$  (**d**), respectively. When not specified,  $C_A = 10^{31}$  cm<sup>6</sup>/s,  $B_r = 10^{-13}$  cm<sup>3</sup>/s, and  $E_t = 0.5$ . In (**b**), we set  $\sigma = 10^{-16}$  cm<sup>2</sup> because we intend to lower the contribution from trap-assisted recombination. The HLI-only TRPL data is used in (**b**) because Auger recombination has negligible effect on TRPL under the LLI condition (Fig. 3a in the main text). When not specified, both HLI and LLI data are used and  $\sigma = 10^{-11}$  cm<sup>2</sup>.



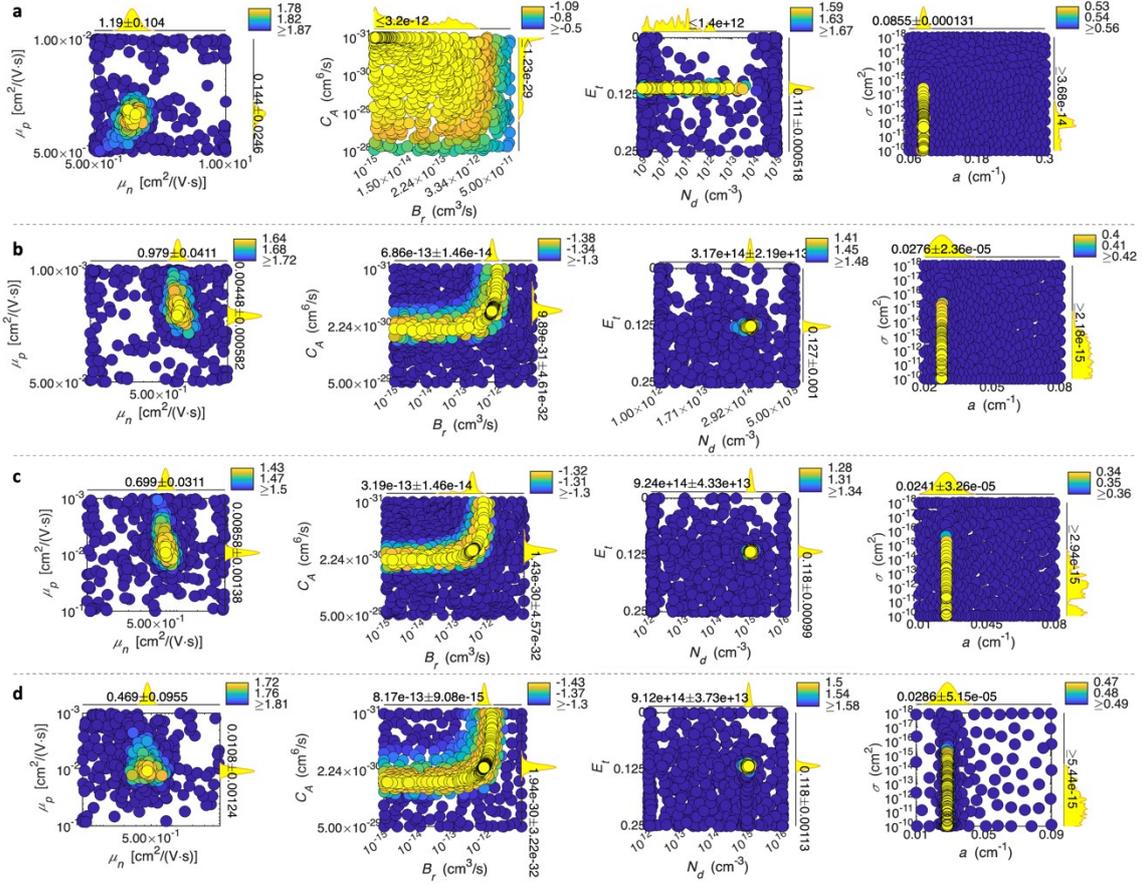
**Fig. S5 | Step 2 of 8D parameter extraction (assuming  $\mu_n > \mu_p$ ) from previously simulated TRPL by our physics-based strategy.** **a**, Extraction of  $\mu_n$  and  $\mu_p$  by setting all parameters as free variables in smaller ranges suggested by Step 1 (Fig. S4). **b**, Extraction of  $B_r$  and  $C_A$  using the HLI-only data and fixed values of  $\mu_n$  and  $\mu_p$  from (a). **c**, Extraction of  $E_t$  and  $N_d$  using fixed values of  $\mu_n$ ,  $\mu_p$ ,  $B_r$ , and  $C_A$  from (a) and (b). **d**, Extraction of  $\sigma$  and  $a$  using the LLI-only data and fixed values of  $\mu_n$ ,  $\mu_p$ ,  $B_r$ ,  $C_A$ ,  $E_t$ , and  $N_d$  from (a)-(c). When not specified, both HLI and LLI data are used. Note that, due to the small effect of  $\sigma$  on the trap-assisted recombination rate [Eq. (14)-(15) in the main text], the extracted  $\sigma$  can show a large uncertainty.



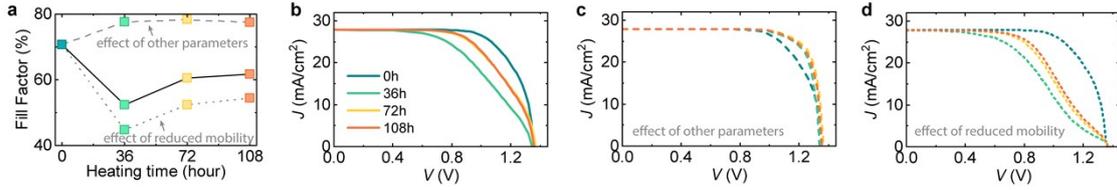
**Fig. S6 | Step 2 of 8D parameter extraction (assuming  $\mu_n < \mu_p$ ) from previously simulated TRPL by our physics-based strategy.** a-d, We use the same strategy as Fig. S5, except that  $\mu_n < \mu_p$  is assumed. The lowest costs obtained at the final steps (c) and (d) are -0.71 and -1.17 respectively, which are higher than that obtained in Fig. S5c-d (-1.31 and -1.51). Therefore, we take the results obtained in Fig. S5 as the final extracted parameters.



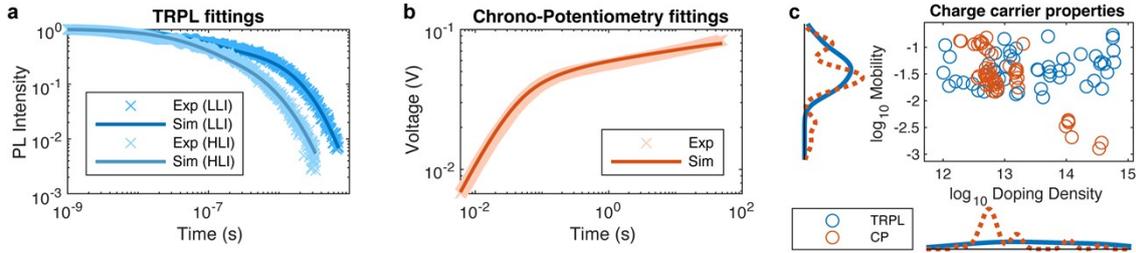
**Fig. S7 | Step 1 of 8D parameter extraction from TRPL experiments of perovskite materials by our physics-based strategy. a-d,** The perovskite materials are placed in nitrogen and heated at 85 °C for different periods of time – 0h (a), 36h (b), 72h (c), and 108h (d). We use the same strategy as Fig. S4. Here we show only the estimated parameters that will be used in Step 2 for simplicity, i.e., the parameters shown in the orange boxes in Fig. S4.



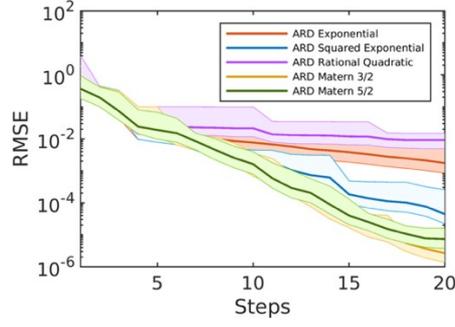
**Fig. S8 | Step 2 of 8D parameter extraction (assuming  $\mu_n > \mu_p$ ) from TRPL experiments of perovskite materials by our physics-based strategy.** a-d, The perovskite material heated for 0h (a), 36h (b), 72h (c), and 108h (d). Here we show only the extracted parameters for simplicity, i.e., the parameters shown in the purple boxes in Fig. S5. When assuming  $\mu_n < \mu_p$ , the extracted parameters are nearly the same with similar costs. This indicates that the extracted mobilities in the left column could have symmetric values (as shown in Fig. S3), e.g., it is possible to have either ( $\mu_n=1.19, \mu_p=0.144$ ) or ( $\mu_n=0.144, \mu_p=1.19$ ) for a sample heated for 0h.



**Fig. S9 | Effect of the material change in perovskite on solar cell performance during thermal degradation by simulation.** **a**, Changes in the fill factor of the solar cells using thermally treated perovskites. The solid line indicates that all parameters of perovskite are taken from Fig. 5e-h of the main text. The dashed line indicates that the mobility values for all samples are the same as the 0h sample, but all other parameters are taken from Fig. 5e, g, and h. In contrast, the dotted line indicates that only the mobility values are taken from Fig. 5f, but all other parameters are the same as the 0h sample. **b-d**, The corresponding current density-voltage ( $J$ - $V$ ) simulation for the coloured squares on the solid line (**b**), dashed line (**c**), and dotted line (**d**) in (**a**). In the  $J$ - $V$  simulation of perovskite solar cells, we consider an ideal structure that consists of electron transport layer, perovskite, and hole transport layer only. The material parameters are specified in Table S2, which are taken from ref<sup>1</sup>.



**Fig. S10 | Using our platform to extract the material parameters of perovskite from two different experiments.** **a,b**, Fittings to TRPL (**a**) and chrono-potentiometry (**b**) experiments of a perovskite by our platform. Chrono-potentiometry is an electrical measurement that applies a step current across the sample and measures the transient voltage. Chrono-potentiometry is a name from the field of electrochemistry<sup>2</sup> and is sometimes referred as the Galvanostatic direct current measurement in perovskite research<sup>3</sup>. **c**, Extracted dopant density and carrier mobility for a different perovskite material from TRPL and chrono-potentiometry (CP). Here we only show the top 50 extractions that have low cost for a clear visualization. The dopant density and carrier mobility extracted from CP show a clear correlation, as indicated by the line formed by the orange circles, which induces an uncertainty in the extracted data. This is a result from the nature of electrical measurement where conductivity – a term that is proportional to the product of dopant density and mobility – dominates. The overlap in the extracted parameters from CP and TRPL reduces the uncertainty in the extracted parameters. This not only confirms the flexibility of our platform in quantitative analysis but also suggests an advantage of our platform in interpreting multiple experiments for better parameter extraction.



**Fig. S11 | Convergence of the extraction of  $N_t$  by different kernel functions.** The extraction is repeated 22 times for each kernel function. The cost function  $f_5$  is used in all extractions.

### 3. Supplementary Tables

Table S1 | Fixed parameters of perovskite and incident laser for TRPL simulation.

Parameter	Value
Relative permittivity	64
Thickness	500 nm
Effective density of states for conduction band	$5 \times 10^{18} \text{ cm}^{-3}$
Effective density of states for valence band	$5 \times 10^{18} \text{ cm}^{-3}$
Electron affinity	3.9 eV
Laser pulse energy	22 or 1960 nJ/cm <sup>2</sup>
Excitation wavelength	477 nm
Optical absorption coefficient at the excitation wavelength	$1 \times 10^5 \text{ cm}^{-1}$
Band gap*	1.6186 eV
Electron mobility ( $M_n$ )*	$1 \text{ cm}^2/(\text{V} \cdot \text{s})$
Hole mobility ( $M_p$ )*	$0.01 \text{ cm}^2/(\text{V} \cdot \text{s})$
Dopant density ( $n_d$ )*	$1 \times 10^{14} \text{ cm}^{-3}$
Defect density ( $n_t$ )*	$5 \times 10^{11} \text{ cm}^{-3}$
Carrier capture cross-section by defects ( $\sigma$ )*	$1 \times 10^{13} \text{ cm}^2$
Relative defect energy level ( $e_t$ )*	0.1
Radiative recombination coefficient ( $b_r$ )*	$1 \times 10^{-11} \text{ cm}^3/\text{s}$
Auger recombination coefficient ( $c_a$ )*	$5 \times 10^{-29} \text{ cm}^6/\text{s}$

\*When not being treated as unknown parameters in BO search or not being specified otherwise, these parameters take the values in the table.

Table S2 | Material parameters for electron transport layer and hole transport layer used in solar-cell simulation (Fig. S9).

Parameter	Electron transport layer	Hole transport layer
Relative permittivity	24	3
Band gap	3 eV	2.2 eV
Electron affinity	3.9 eV	3.1 eV
Effective density of states for conduction band	$1 \times 10^{20} \text{ cm}^{-3}$	$1 \times 10^{20} \text{ cm}^{-3}$
Effective density of states for valence band	$1 \times 10^{20} \text{ cm}^{-3}$	$1 \times 10^{20} \text{ cm}^{-3}$
Electron mobility	$2 \text{ cm}^2/(\text{V}\cdot\text{s})$	$0.01 \text{ cm}^2/(\text{V}\cdot\text{s})$
Hole mobility	$0.01 \text{ cm}^2/(\text{V}\cdot\text{s})$	$0.1 \text{ cm}^2/(\text{V}\cdot\text{s})$
Dopant density	$1 \times 10^{17} \text{ cm}^{-3}$ (donor)	$1 \times 10^{17} \text{ cm}^{-3}$ (acceptor)
Thickness	50 nm	50 nm

All parameters in this table are taken from ref<sup>1</sup>.

#### 4. Supplementary References

1. J. Peng, D. Walter, Y. Ren, M. Tebyetekerwa, Y. Wu, T. Duong, Q. Lin, J. Li, T. Lu, M. A. Mahmud, O. L. C. Lem, S. Zhao, W. Liu, Y. Liu, H. Shen, L. Li, F. Kremer, H. T. Nguyen, D.-Y. Choi, K. J. Weber, K. R. Catchpole and T. P. White, *Science*, 2021, **371**, 390.
2. A. J. Bard and L. R. Faulkner, *Electrochemical Methods: Fundamentals and Applications*, John Wiley & Sons, New York, USA, 2000.
3. T.-Y. Yang, G. Gregori, N. Pellet, M. Grätzel and J. Maier, *Angew. Chem. Int. Ed.*, 2015, **54**, 7905-7910.