

Supplementary Information for Machine Learning for Analysis of Time-resolved Luminescence Data

Nikola Đorđević^{*▲}, Joseph S. Beckwith[°], Maksym Yarema^{*}, Olesya Yarema^{*}, Arnulf Rosspeintner[°], Nuri Yazdani^{*},
Juerg Leuthold[▲], Eric Vauthey[°], Vanessa Wood^{*}

^{*}Materials and Device Engineering Group, ETH Zurich, 8092 Zurich, Switzerland

[▲]Institute of Electromagnetic Fields, ETH Zurich, 8092 Zurich, Switzerland

[°]Physical Chemistry Department, University of Geneva, 1211 Geneva, Switzerland

Decay rate analysis of Ag-In-Se nanocrystals

Statistical Analysis of Different Models

Statistical analysis of different *a priori* models used to fit TRPL data for emission from Ag-In-Se nanocrystals (Figure 1a-c in the main text) is summarized in Table S1. The usual goodness-of-fit metrics for fitting using the least squares method are R^2 and R_{adj}^2 statistics (often called coefficients of determination), which are calculated using the following formulas:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2};$$

$$R_{adj}^2 = 1 - (1 - R^2) \frac{n - 1}{n - p - 1},$$
(S1)

where n is the number of data points, p the number of free parameters (degrees of freedom), y_i measured data, \hat{y}_i are estimated values from the model and \bar{y} is the mean value of the measured data. Another metric often reported when analyzing TRPL datasets is the Neumann formula¹⁻² for reduced χ^2 . This is given by: $\chi_N^2 = \frac{1}{n-p} \sum_{i=1}^n \frac{(y_i - \hat{y}_i)^2}{y_i}$, where the meaning of the parameters is the same as described above. The values of χ_N^2 indicate the goodness-of-fit, taking into account degrees of freedom and the fact that each measured data point y_i has different variance $\text{var}(y_i) = \sqrt{y_i}$. If $\chi_N^2 < 1$, the model overfits the data, while in the case $\chi_N^2 > 1$, model underfits the data. Thus, the closer the value of χ_N^2 to unity, the better the model is.

Table S1. Summary of R^2 and χ_N^2 statistics for a priori models used for fitting measured TRPL data for Ag-In-Se nanocrystals. Please note that constant term is not included in degrees of freedom.

Fitting Model	Degrees of freedom (p)	R^2	R_{adj}^2	χ_N^2
Bi-exponential	4	0.9991	0.9991	6.5
Tri-exponential	6	0.9996	0.9996	2.7
Mixture model	5	0.9996	0.9996	2.5

Calculated Decay Rate Distribution Using Delta-Basis

The calculated decay rate distribution of measured Ag-In-Se nanocrystals using delta-basis is shown in Figure S1. The resulting distribution of decay rates features a narrow distribution that is centered around $6 \times 10^{-3} \text{ ns}^{-1}$, and an additional broad feature that is centered at $\sim 2.6 \times 10^{-2} \text{ ns}^{-1}$, again confirming that the mixture model (i.e., single exponential + stretched exponential) is an appropriate choice for fitting TRPL of these nanocrystals. In other words, the conclusion drawn from the machine learning is independent of the chosen basis: the histogram basis in the main text and the delta-function basis both point toward the same model, although the results may be clearer in some bases than others.

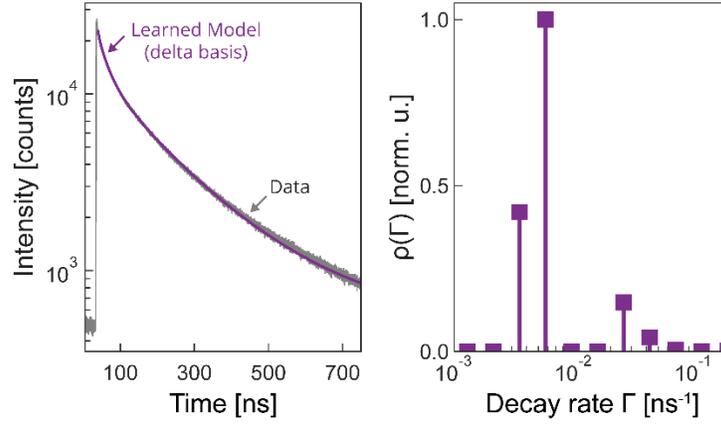


Figure S1. Time-resolved photoluminescence of Ag-In-Se nanocrystals (same as in Figure 1 in the main text). Left: The model obtained using delta-basis (purple) along with the measured data (gray). Right: The calculated decay rate distribution, indicating a narrow distribution centered around $6 \times 10^{-3} \text{ ns}^{-1}$ and a broader distribution centered around $2.6 \times 10^{-2} \text{ ns}^{-1}$.

Simulated Datasets

Biexponential Decay

To simulate the biexponential decay, we set the analytic decay

$$y_{an}(t) = A_1 \exp(-\Gamma_1 t) + A_2 \exp(-\Gamma_2 t) \quad (\text{S2})$$

and let $A_1 = A_2$. To have proper understanding of physical process driving the emission, one has to take into account proper normalization³ of the amplitudes A_i . The fitting amplitudes A_i contain the information about both radiative and nonradiative decay³. For this analysis, we assume that two channels have a unique radiative rate and that decay rate distribution is simply weighted with the quantum yield of a channel and the total decay rate. That is why the actual decay rate distribution shown in Fig. 3 (main text) is normalized in such a way that

$$\rho(\Gamma_i) = A_i / \Gamma_i.$$

To emulate real-world experiment, we add noise to the analytical datasets. The main sources of noise when performing photon-counting experiments is background noise (having a constant value) and the counting noise, which stems from Poisson distribution of photon counts⁴. Since the Poisson noise is additive, the final signal that is used reads:

$$y(t) = PN[y_{an}(t) + noise], \quad (\text{S3})$$

where $PN[x]$ represents a random number following Poisson distribution with the mean value of x . We set the decay amplitude in such a way that the peak count value (at $t=0$) is $\sim 10^3$ times higher than the background noise.

Stretched-exponential Decay

To validate our approach in the situation when the decay rate distribution is broad and asymmetric, we analyze the analytical decay following the Kohlrausch (stretched-exponential) decay function⁵

$$y_{an}(t) = A \exp\left(-\left(\frac{t}{\tau_{ww}}\right)^\kappa\right). \quad (\text{S4})$$

For this type of decay, there is an analytic form of the decay rate distribution

$$\rho_\kappa(\Gamma) = \frac{\tau_{ww}}{\pi} \sum_{n=1}^{\infty} (-1)^{n+1} \frac{(\Gamma \cdot \tau_{ww})^{-1+n\kappa}}{n!} \tilde{\Gamma}(1+n\kappa) \sin n\kappa\pi, \quad (\text{S5})$$

where $\tilde{\Gamma}$ is the gamma-function. In our analysis we set $\tau_{WW} = 25$ ns and $\kappa = 0.6$. The Poisson and the background noise are added to the analytical decay function (S4) in the same way as discussed for the biexponential function.

Log-normal Distribution of Decay Rates

A log-normal distribution of decay rates is defined as

$$\rho(\Gamma) = \frac{0.5}{\Gamma \cdot \sigma \sqrt{2\pi}} \exp\left(-0.5 \left(\frac{\log \Gamma - \mu}{\sigma}\right)^2\right) \quad (S6)$$

Bi-modal Log-normal Distribution of Decay Rates

In addition to the stretched distribution of decay rates presented in the main text and given in (S5), we analyze a bi-modal log-normal decay rate distribution, with the following form:

$$\rho(\Gamma) = \frac{0.5}{\Gamma \cdot \sigma_1 \sqrt{2\pi}} \exp\left(-0.5 \left(\frac{\log \Gamma - \mu_1}{\sigma_1}\right)^2\right) + \frac{0.5}{\Gamma \cdot \sigma_2 \sqrt{2\pi}} \exp\left(-0.5 \left(\frac{\log \Gamma - \mu_2}{\sigma_2}\right)^2\right). \quad (S7)$$

Here, we define the distribution parameters $(\mu_1, \sigma_1, \mu_2, \sigma_2)$ such that the two centers of distribution are located at $\Gamma_{m1} = \exp(\mu_1 - \sigma_1^2) = 8 \times 10^{-3}$ ns⁻¹ and $\Gamma_{m2} = \exp(\mu_2 - \sigma_2^2) = 0.2$ ns⁻¹. We choose widths of the distributions to be $\sigma_1 = 0.6$ and $\sigma_2 = 0.2$, so that the mixture of one narrow and one wider distribution of decay rates is simulated. Poisson noise is added to the data in the same way as explained above, and we set a signal-to-noise ratio (SNR) of 100, as explained in the section about the resolution limits of our approach (below).

Calculated decay rate distributions using delta- and histogram-basis based on such a simulated decay are shown in Figure S2. As can be seen in the figure, both the delta- and the histogram-basis calculations correctly predict two decay distributions with different widths. The calculated distribution using delta-basis is able to ‘track’ the shape of this bi-modal distribution, but apart from the shape of the broad distribution, it does not sample the peak value properly. The histogram-basis on the other hand correctly traces out the wider distribution, and correctly predicts the narrow one. This example highlights the importance of trying out the different basis sets in LumiML to improve understanding of the underlying decay rate distribution.

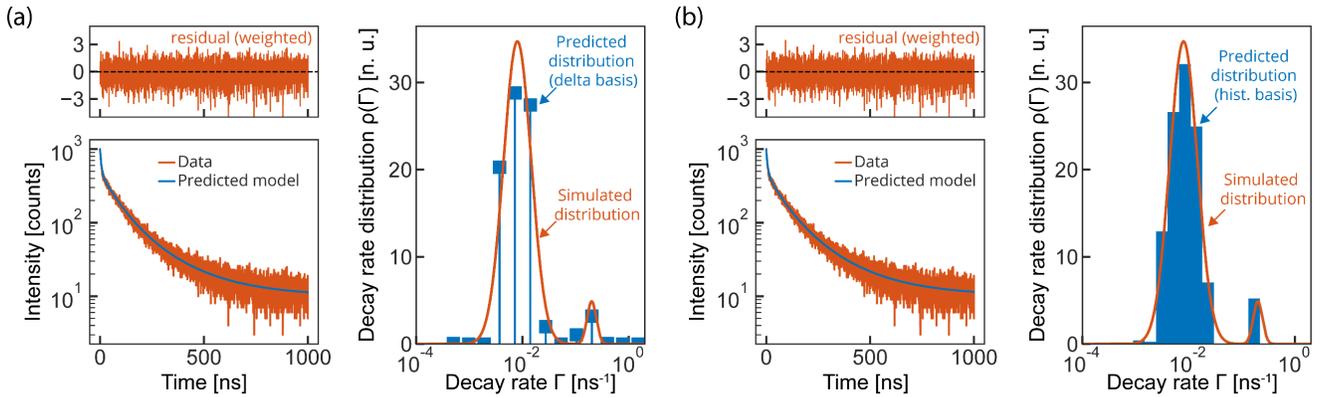


Figure S2. Calculated decay rate distribution for simulated bi-modal log-normal distribution of decay rates. (a) Left: generated data and model prediction (bottom) and weighted residuals (top); right: Simulated decay rate distribution and calculated decay rate distribution using delta-basis. (b) Left: generated data and model prediction (bottom) and weighted residuals (top); right: Simulated decay rate distribution and calculated decay rate distribution using histogram-basis.

Impact of the Regularization Parameters

Effect of the Balance Between L1 and L2 penalties (parameter θ)

The value of the parameter θ controls the ratio between L1 and L2 penalties. In the case $\theta = 0$, only the L2 penalty (i.e. Ridge regression) is applied; with $\theta = 1$, only the L1 penalty is present (i.e., Lasso regression). In the case of the Lasso penalty, once the coefficients in the sought vector β are correlated, the Lasso randomly picks which coefficients will be shrunk. This could lead to wrong interpretation of a calculated decay rate distribution.

Here, we explore this impact by setting the parameter θ and evaluating its effect on reconstruction of the decay rate distribution. The simulated distribution of decay rates is a log-normal distribution defined in (S6) and has its center at $\Gamma_m = \exp(\mu - \sigma^2) = 0.05 \text{ ns}^{-1}$. We analyze two cases, broad distribution of decay rates having $\sigma = 0.7$, and a narrower distribution having $\sigma = 0.1$. Poisson noise is added to the data in the same way as explained in the sections above, having a signal-to-noise ratio of 100 (please see the section Resolution Limits for Recovery of Decay Rate Distribution). To evaluate the performance of different values of θ , we divide the simulated data set into a training set having $\sim 2/3$ of all the datapoints, and a validation set, on which we evaluate the performance (log-likelihood loss defined in Eq. (5) in the main text). We use the same training set for all different θ values. For each of the values chosen for θ , we perform 5-fold cross-validation to obtain the optimal α , and evaluate the performance (log-likelihood loss) on the hold-out (validation) set.

Resulting decay rate distributions for different values of θ using a histogram-basis are shown in Figure S3 and Figure S4, for broader and narrower decay rate distribution, respectively. Additionally, in Figure S4, we show the ‘average strength’ of the decay rate distribution within each ‘bin’ used in histogram basis. We calculate the ‘average strength’ within bin spanning from Γ_i to Γ_{i+1} as $\psi_i = \int_{\Gamma_i}^{\Gamma_{i+1}} d\Gamma \rho(\Gamma) / (\Gamma_{i+1} - \Gamma_i)$. Due to the way the histogram basis is defined, this ‘average strength’ of simulated decay rate distribution allows us to compare contribution of calculated distribution within each histogram bin to the decay with the analytical case (simulated distribution).

The resulting log-likelihood loss is summarized in the

Table S2. From

Table S2, it can be concluded that the value of $\theta = 0.5$ gives the best performance for broad distribution of decay rates, although there is no significant visual difference in Figure S3 for the reconstructed decay rate distributions up to $\theta = 0.7$. In the case of narrow distribution of decay rates, the best performance is achieved for setting $\theta = 1$, i.e. using Lasso regression, which yields sparse solution for the β vector.

These results emphasize that appropriate choice of hyperparameter θ is also data-dependent. For example, in the case of a dataset with high signal-to-noise ratio, which shows strong non-monoexponential behavior, a good initial guess would be to start with $\theta = 0.5$. However, if one tries to recover a narrow distribution of decay rates, and the resolution of the decay rate axis is not sufficiently high due to the noise (please see the section about the resolution limits) or the dataset contains < 1000 data points, one should use values of θ that are closer to 1 since the Lasso regularization can yield sparse solutions.

Table S2. Summary of log-likelihood loss for different values of θ .

<i>Value of θ</i>	Log-likelihood Loss (validation set)	
	Broad Distribution ($\sigma = 0.7$)	Narrow Distribution ($\sigma = 0.1$)
<i>0</i>	-120.566016	-160.5334
<i>0.3</i>	-120.5660505	-160.5739
<i>0.5</i>	-120.5662786	-160.5989
<i>0.7</i>	-120.5661291	-160.6232
<i>1</i>	-120.5654109	-160.6581

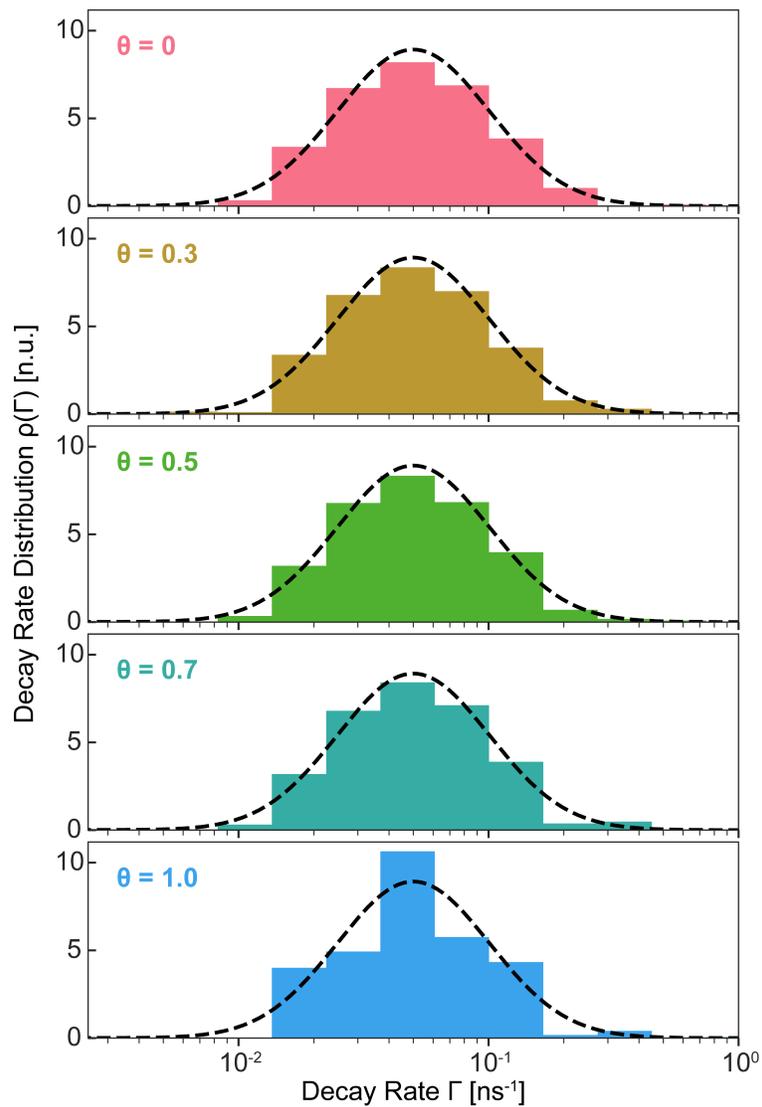


Figure S3. Resulting decay rate distribution for different values of θ ranging from 0 (Ridge regression) towards 1 (Lasso regression) using histogram-basis. The simulated broad log-normal distribution is shown in dashed line. Although up to $\theta = 0.7$, there is no prominent visual difference in the reconstructed decay rate distributions, log-likelihood loss (Eq (5) in the main text) is minimized for $\theta = 0.5$. The bottom figure ($\theta = 1$) illustrates the problem of Lasso regression when the ground truth distribution is wide. The Lasso choses the coefficient closest to the distribution peak and shrinks all other coefficients, thus the resulting distribution appears narrower than it really is.

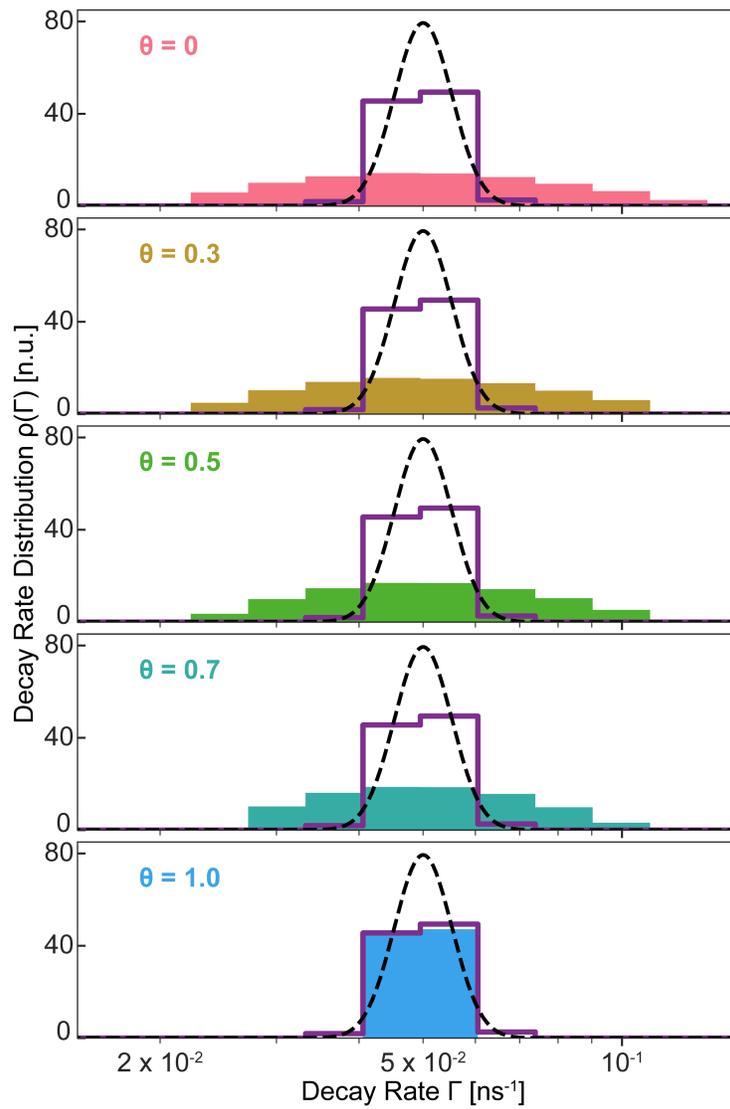


Figure S4. Resulting decay rate distribution for different values of θ ranging from 0 (Ridge regression) towards 1 (Lasso regression) using histogram-basis. The simulated narrow log-normal distribution is shown in dashed line. The purple solid line represents 'average strength' of the simulated decay rate distribution as explained in the text. In the case of narrow underlying distribution of decay rates, the calculated coefficients best describe the underlying distribution for the case of $\theta = 1$.

Impact of amount of regularization (parameter α)

To highlight the importance and the effect of the regularization parameter α on values of the sought coefficients β and thus on the calculated decay rate distribution, we calculate the β for different values of the α parameter while keeping the ratio of the L1 and L2 penalties balanced (i.e. $\theta = 0.5$). Here, we analyze the simulated log-norm distribution defined in (S6), and use the same distribution parameters as well as the same training/validation sets division strategy as in the section above.

Large values of α (large penalty terms) shrink the L2 norm of the solution β towards zero. This effect is illustrated in Figure S5. The prediction loss shown in the Figure S5 is evaluated on the validation set. Prediction loss (red line) is smallest when α is small. Comparing the prediction loss values in Figure S4 to those the table 2, shows that, in the remine of small alpha, the effect of varying θ from 0 to 1 is to slightly change the prediction loss (i.e., for all values of θ , the prediction loss is around -120.56)

Resulting decay rate distribution for the different α are shown in the Figure S6. Large regularization parameters drive all of the coefficients β towards zero, resulting in completely wrong prediction. On the other hand, too small an α (almost no regularization at all) leads towards a noisy reconstruction of the decay rate distribution, and is thus not reliable.

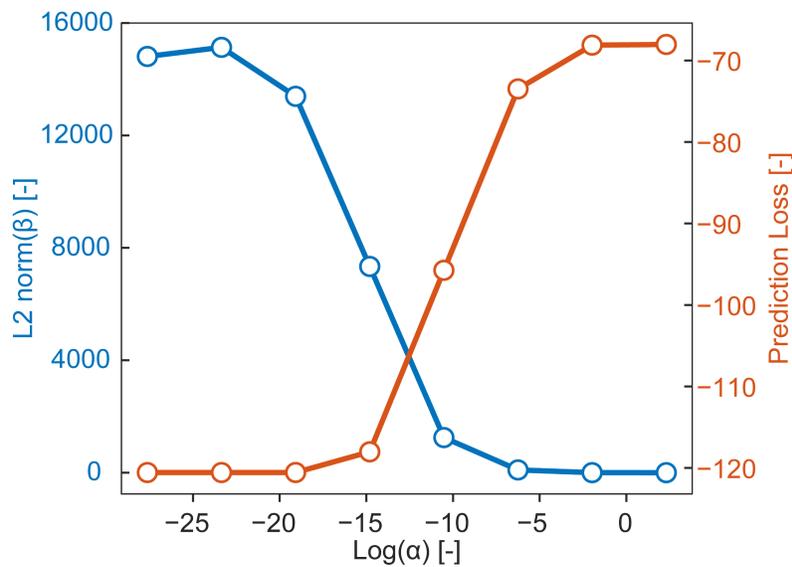


Figure S5. Dependence of the L2 norm of the calculated coefficients β and prediction loss (log-likelihood loss defined in Eq. (5) in the main text) on the regularization parameter α . The prediction loss is evaluated on the validation set, while the L2 norm is evaluated on the training set.

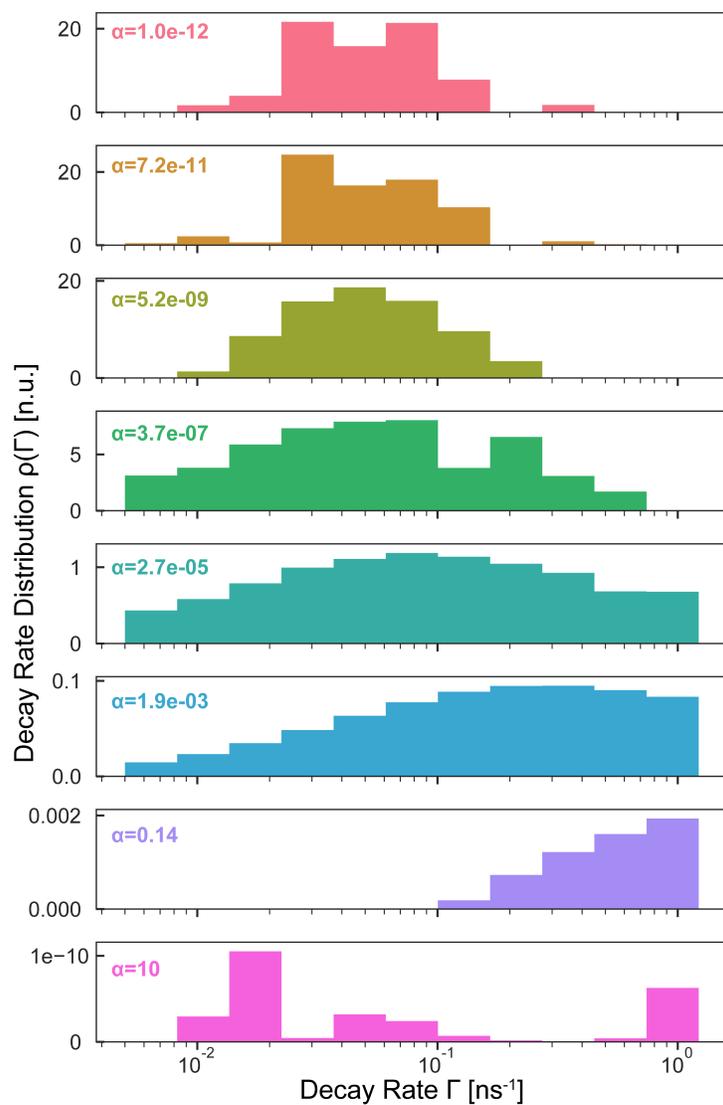


Figure S6. Calculated decay rate distribution for different amount of regularization controlled by parameter α . As discussed in the main text, a large value of α causes all coefficients β to be close to zero (note y-axis scale in the bottom graph where $\alpha = 10$), leading to a completely wrong prediction of the distribution of decay rates.

Formulation of the Inverse Problem

As shown in the main text, the elements of the kernel matrix \mathbf{X} are obtained by evaluating Eq (1) from the main text across each basis function. Although the triangular and histogram basis do not require post-normalization of coefficients, this is not the case with delta-basis⁶⁻⁹. The formulation of the kernel matrix is summarized in Table S3.

Table S3. Summary of matrix problem formulation. (Eq. (4) in the main text)

Basis	Basis function $\varphi_i(\Gamma)$	Matrix Element x_{ij}	Normalization
Delta	$\delta(\Gamma - \Gamma_i)$	$\exp(-\Gamma_i t_j)$	Yes ^{6, 9}
Histogram	$\theta_h(\Gamma - \Gamma_i) - \theta_h(\Gamma - \Gamma_{i+1})$	$\frac{1}{t_j} (\exp(-\Gamma_i t_j) - \exp(-\Gamma_{i+1} t_j))$	None
Triangular	$\chi_i(\Gamma) = \begin{cases} \frac{\Gamma - \Gamma_{i-1}}{\Gamma_i - \Gamma_{i-1}}, \Gamma \in (\Gamma_{i-1}, \Gamma_i] \\ \frac{\Gamma_{i+1} - \Gamma}{\Gamma_{i+1} - \Gamma_i}, \Gamma \in (\Gamma_i, \Gamma_{i+1}] \end{cases}$	$\frac{e^{-t_j(\Gamma_{i-1} + \Gamma_i)} (-e^{t_j \Gamma_i} + e^{t_j \Gamma_{i-1}} (1 - t_j \Gamma_{i-1} + t_j \Gamma_i))}{t_j^2 (\Gamma_{i-1} - \Gamma_i)} + \frac{e^{-t_j(\Gamma_i + \Gamma_{i+1})} (-e^{t_j \Gamma_i} + e^{t_j \Gamma_{i+1}} (1 + t_j \Gamma_i - t_j \Gamma_{i+1}))}{t_j^2 (\Gamma_i - \Gamma_{i+1})}$	None

Resolution Limits for Recovery of Decay Rate Distribution

As mentioned in the main text, the decay rate axis is discretized using the exponential sampling method (ESM). ESM puts the individual decay rates used for analysis equidistant on a logarithmic scale. The spacing between individual rates is given by

$$\Gamma_k = \Gamma_0 \exp(k\pi/\omega), \quad (\text{S8})$$

where ω is resolution parameter. Since it is difficult (or impossible) to know beforehand the optimal resolution one should use for the analysis, we cannot assume a single set of resolution limits that are valid for all possible datasets.

However, here we show how the background noise influences the decay rate distribution calculated from computer-generated datasets using different resolution parameters. From previous studies⁸⁻⁹, it is clear that the maximal resolution allowed will depend on the noise in the data. Fletcher and Ramsay⁷⁻⁸ studied this problem for optimal choice of parameters for the histogram-basis, concluding that it is beneficial to artificially increase resolution by stacking and shifting multiple low-resolution base histograms. They, however, added white-noise to the entire signal, so we cannot strictly follow their arguments.

Thus, to achieve some practical guidelines towards choosing the resolution of decay rate distribution calculation, we generate the TRPL datasets that follow the log-normal distribution of decay rates $\rho(\Gamma) = \frac{1}{\Gamma \cdot \sigma \sqrt{2\pi}} \exp\left(-0.5 \left(\frac{\log \Gamma - \mu}{\sigma}\right)^2\right)$. In our case, we define log-normal distribution parameters such that it has center at $\Gamma_m = \exp(\mu - \sigma^2) = 0.025 \text{ ns}^{-1}$, and $\sigma = 0.5$. The analytical data are given by evaluating Eq. (1) in the main text. To study the influence of the noise on the decay rate distribution, we set the background noise constant and change the ratio of the maximum number of counts to the background noise. That way, we define signal-to-noise ratio (SNR) as $SNR = \frac{y_{max}}{\langle background \rangle}$, and study SNRs from 2-100, which is within realistic bounds. This relates to the real-life measurements in the following way: we let the photon counting unit collect the data for exactly the same time, hence the same background counts, so higher values of SNR mean that an artificial emitter is more efficient, therefore giving SNR times more peak counts than the background average. The simulated datasets are shown in Figure S7.

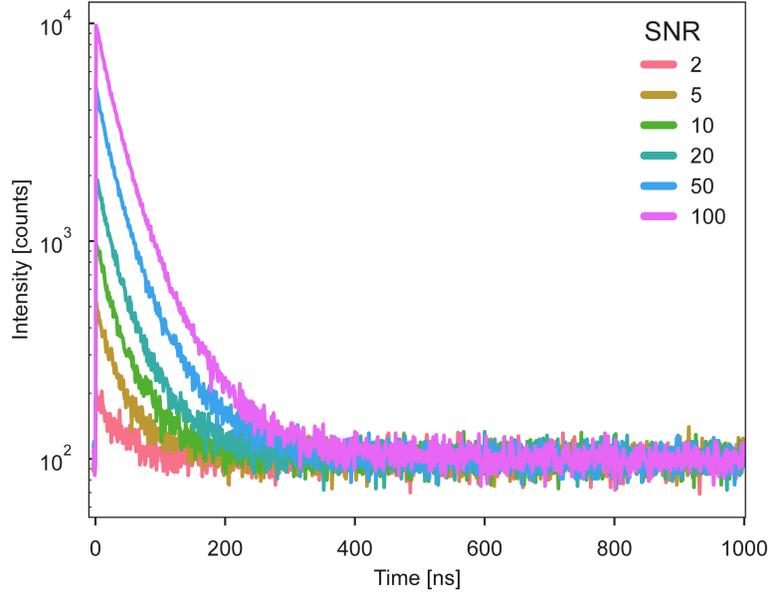


Figure S7. Simulated TRPL datasets.

Figure S8 shows the calculated decay rate distribution for the above-mentioned datasets with resolution parameter $\omega = \pi - 5\pi$, with steps of 0.5π using the histogram basis. It is clear from Figure S8 that the low-resolution case $\omega = \pi$ gives important information about the decay. Namely, the results show that there is a broad distribution of decay rates. Another important conclusion is that the same information is obtained across the entire range of SNRs. This is one important guideline: low resolution scan can give the information about whether the decay rate distribution is broad. However, if one tries to obtain finely resolved information about the distribution, the noise in the signal becomes very important. As can be seen, for all resolution parameters $\omega > \pi$ taken into account here and signals with SNRs 2-10, the calculated decay rate distributions do not match the original one. For example, in the case of SNR=10 and $\omega = 3.5\pi$, the calculated distribution correctly predicts broad distribution centered $\sim 0.025 \text{ ns}^{-1}$, but also predicts a spurious decay component at $\sim 10^{-3} \text{ ns}^{-1}$. However, as signal-to-noise ratio reaches values of 20, which can be regularly achieved with combination of reasonably good equipment and good emitters, all resolution parameters studied here correctly predict the simulated decay rate distribution. Important to note here is that with increasing SNR as we defined it here, we can obtain better information about the slowly decaying components. Therefore, i.e. SNR=20 and $\omega = 3.5\pi$ has apparent ‘cut-off’ at $\sim 10^{-2} \text{ ns}^{-1}$, while in SNR=50 and $\omega = 3.5\pi$ case one can resolve slower components as well.

Figure S9 shows calculated decay rate distribution for above mentioned datasets with resolution parameter $\omega = \pi - 4\pi$, with steps of 0.5π using the delta basis. The most notable difference comparing to the histogram basis is proper sampling of the simulated decay rate distribution for low resolution ($\omega = \pi$) case, even for the low SNR of two. However, for this specific comparison, one can see that the delta basis properly reconstructs the decay rate distribution only up to $\omega = 3\pi$ even for SNR = 100. As we keep increasing the resolution, artificial ripples⁹ are observable in the low decay rate range (please see (SNR=100, $\omega = 3.5\pi$) case in Figure S9).

We also calculated the decay rate distribution using a triangular basis (Figure S10) for the SNR = 2-100 and resolution parameters $\omega = \pi - 4\pi$. The resulting distribution describes the simulated log-normal distribution well for resolution ω up to 2π , across all simulated signal-to-noise ratios. However, some instabilities are visible if the resolution increases, which is probably due to numerical optimization that is used to obtain the coefficients for reconstruction of the decay rate distribution. This is apparent if we look at the calculated distribution for the SNR=100 and resolution parameters of 3π and i.e. $\omega = 3.5\pi$. In this case, lower resolution case shows large ripples that are not physical, while by increasing the resolution, those ripples disappear. This occasional instability can come from Nelder-Mead algorithm that is used to solve optimization problem in Eq. (7) in the main text and will be subject of subsequent analysis and work that we plan to perform.

The main guidelines that we want to point out from this case-study is that for low SNR situations, one should use all three basis sets (and low resolution) to confirm whether the underlying distribution of decay rates is broad and/or if there are more

decay channels that are driving the luminescence. In the range of SNRs of 5 – 20, resolution of up to $\omega = 2\pi$ could be used, while for high signal-to-noise ratio measurements, histogram basis seems particularly appealing, since it is able to reconstruct underlying distribution of decay rates with high resolutions, up to $\omega = 5\pi$ in this particular case. These guidelines are, however, based upon the analysis presented here, so they should be taken as a rule-of-thumb rather than strict how-to approach, since the results will be largely data-dependent.

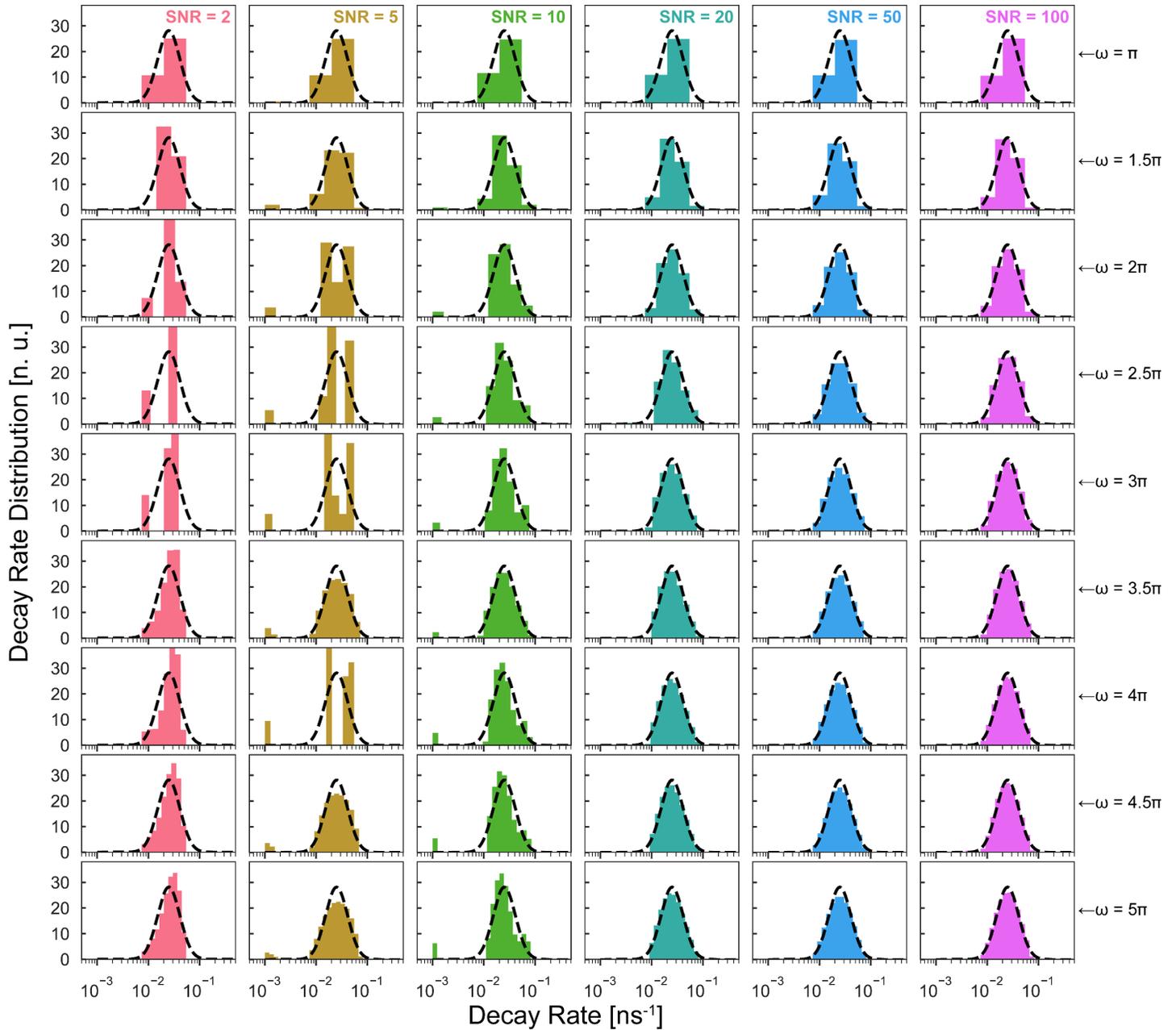


Figure S8. Calculated decay rate distribution using histogram basis for datasets shown in Figure S7. The SNR is increased along the rows from 2 – 100 (color-coded as well), while resolution parameter is increased from top ($\omega = \pi$) to bottom ($\omega = 5\pi$) in each column. Dashed line represents simulated log-normal distribution (please see text for details).

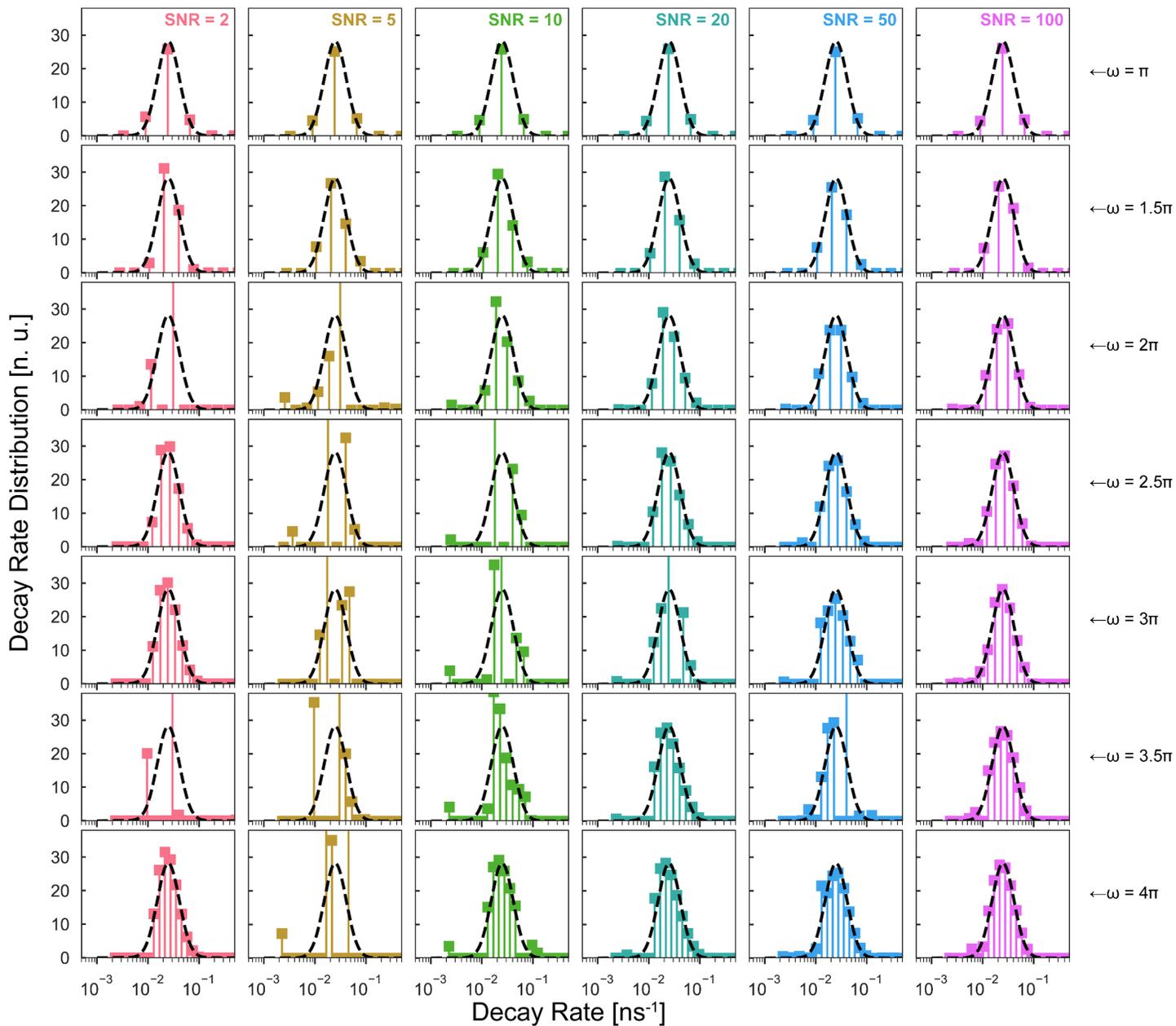


Figure S9. Calculated decay rate distribution using delta basis for datasets shown in Figure S7. The SNR is increased along the rows from 2 – 100, while resolution parameter is increased from top ($\omega = \pi$) to bottom ($\omega = 4\pi$) in each column. Dashed line represents simulated log-normal distribution (please see text for details).

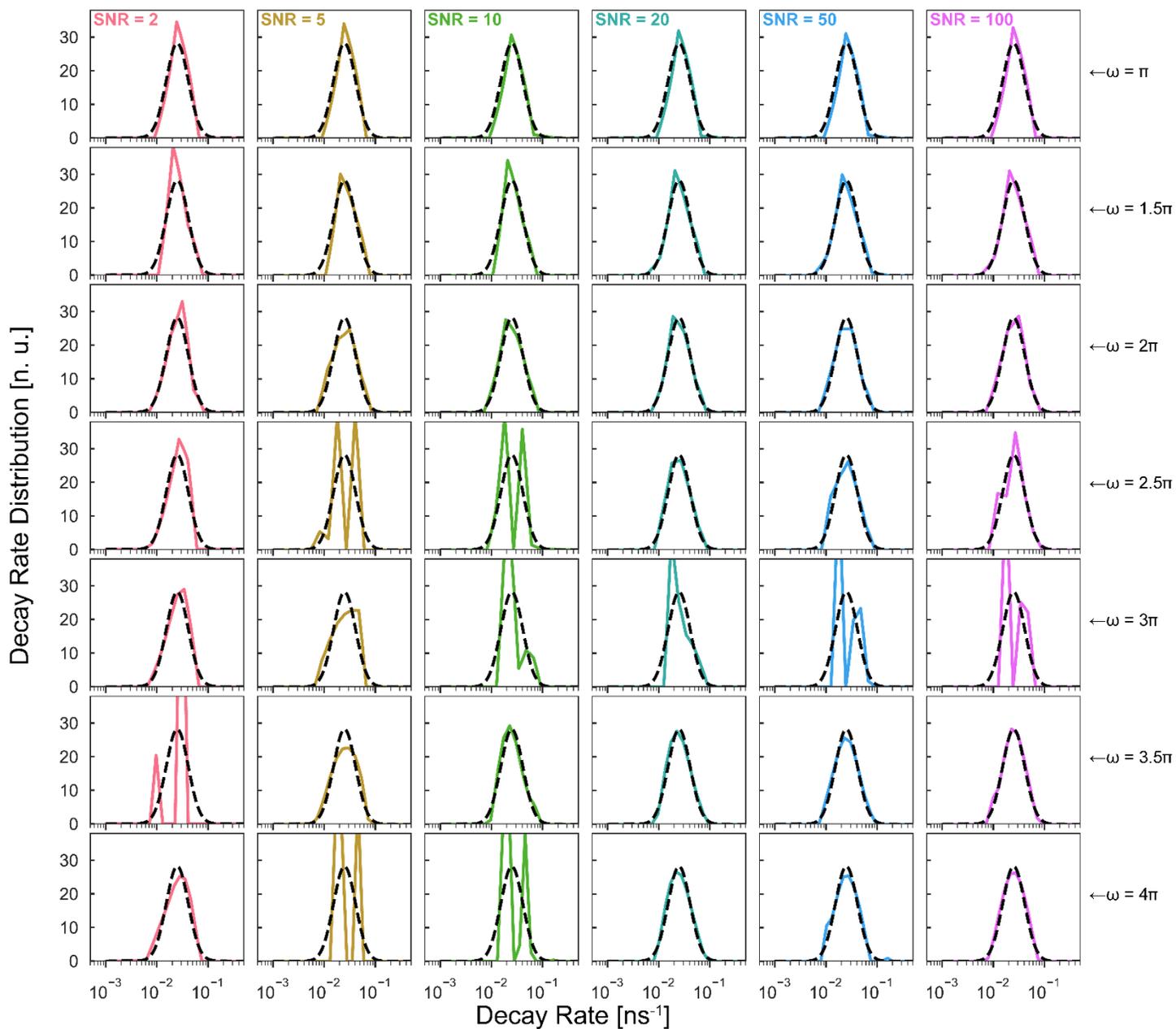


Figure S10. Calculated decay rate distribution using triangular basis for datasets shown in Figure S7. The SNR is increased along the rows from 2 – 100, while resolution parameter is increased from top ($\omega = \pi$) to bottom ($\omega = 4\pi$) in each column. Dashed line represents simulated log-normal distribution (please see text for details)

CsPbBr₃ Nanocrystals

Steady-state absorption, integrated photoluminescence and high-resolution transmission electron microscopy (HRTEM) image of in-house synthesized CsPbBr₃ nanocrystals¹⁰⁻¹¹ are shown in Figure S11.

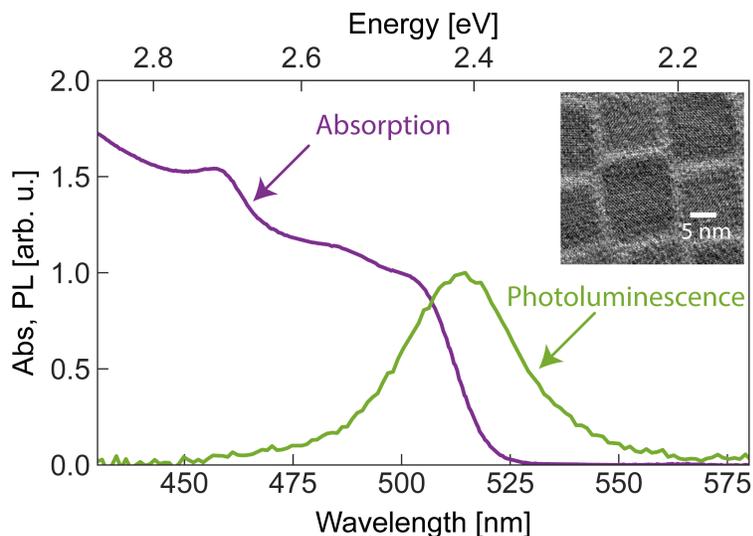


Figure S11. Absorption (purple) and photoluminescence (green) spectra of CsPbBr₃ nanocrystals (inset: HRTEM image of nanocrystals showing cubic shape with average cube length of 9.7 nm).

Ultrafast Spectroscopy of CsPbBr₃ Nanocrystals

FLUPS Maps

Figure S12 shows the measured energy- and time-resolved PL spectra for different pump pulse energies ranging from 5 – 100 nJ (0.13 – 2.6 mJ/cm²). Please note that the scale on the time axis is linear up to 2 ps, and then switches to a logarithmic one.

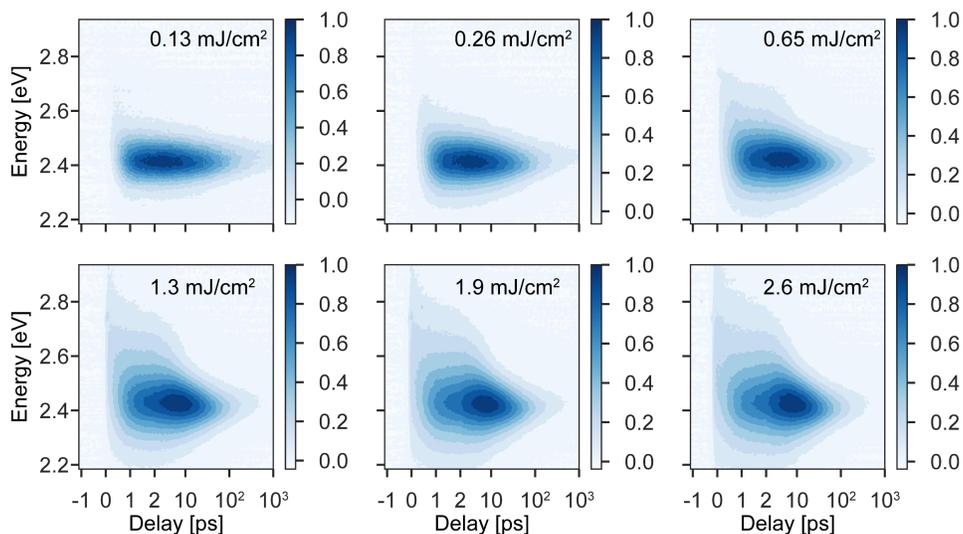


Figure S12. Two-dimensional contour plot of time- and energy-resolved emission intensity from FLUPS experiment with varying pump fluence (insets). All plots are normalized to the maximal measured value.

It is evident from the Figure S12 that with increasing the pump fluence, the photoluminescence spectrum gets substantially broadened to the high-energy side. This broadening indicates population of higher-energy states above the band edge, which rapidly relax within couple of picoseconds. Emission of those states is visualized in Figure S13, where PL spectra taken at different time delays from 0.1 – 300 ps are normalized to unit area. As noted above, for low pump fluence, the signal of states above the band-gap is within the noise, while it becomes apparent for pump fluences above 0.26 mJ/cm².

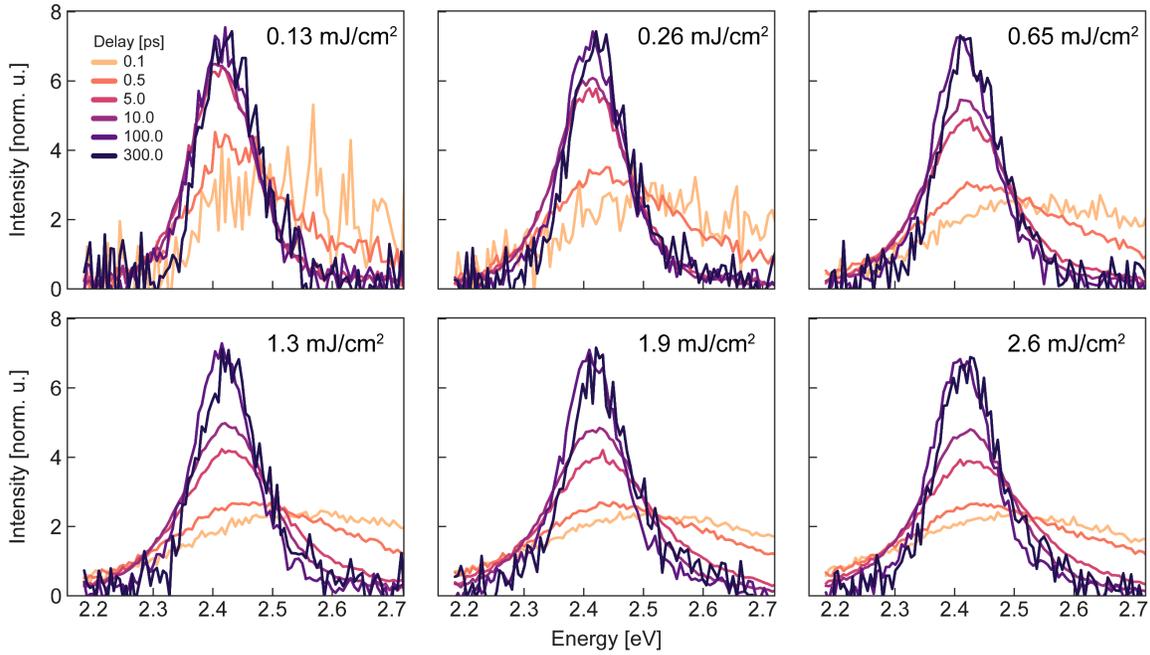


Figure S13. Photoluminescence spectra taken at different time delays (color coded). Each spectrum is normalized to unit area under the curve. Legend is identical for all subplots.

In addition, one can see that in time, position of the PL intensity peak shifts from ~ 2 ps in case of pump fluence of 0.13 mJ/cm^2 , towards ~ 10 ps for the highest fluence used. This is another strong indication of carrier cooling. Since the pulse width (temporal) of the pump laser is ~ 100 fs, the build-up of the band edge emission that takes 2 ps should originate from the carriers that rapidly cool down to the states at the band gap extrema. Therefore, by measuring how long does it take for band-edge PL to reach its maximum value, we can give an estimate of the carrier relaxation time. If we assume that the carrier cooling is much faster than the rest of the emission processes, we can approximate the time constant for cooling as

$$\tau_{cool} \approx t_{build-up}/5. \quad (S9)$$

This approximation is based on the fact that assuming only carrier cooling process, the build up of the band edge signal would take the form of $I(t) \sim (1 - \exp(-t/\tau_{cool}))$, and thus will reach $\sim 99\%$ of the maximal value after $t \approx 5 \cdot \tau_{cool}$. The estimation of the carrier relaxation time from (S9) is shown in Figure S14. In the case of low pulse energy of 5 nJ (low fluence, close to the condition of one exciton per NC on average), we extract the relaxation time of ~ 0.35 ps, which is close to the measured value of ~ 0.31 ps from transient absorption measurements¹², also performed in the low pump fluence regime.

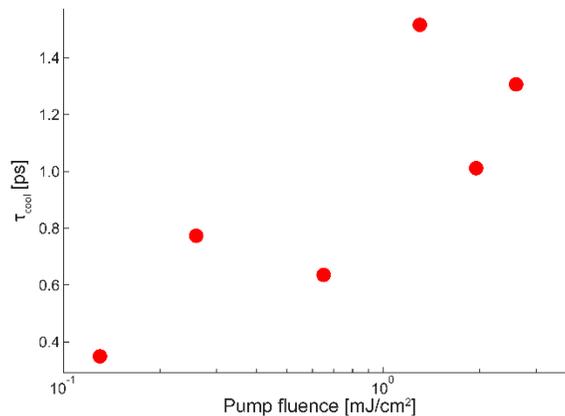


Figure S14. Estimated carrier relaxation time for different pump fluences, which is a direct measure of the number of excited carriers.

Integrated PL

Figure S15 shows the spectra shown in Figure S12, integrated along time axis and normalized such that peak value equals to one. The broadening of the low-energy portion of the spectrum can be explained by formation of charged and multi-excitons in the nanocrystals, while the increase of the high energy side can be explained by emission from the states above the band edge.

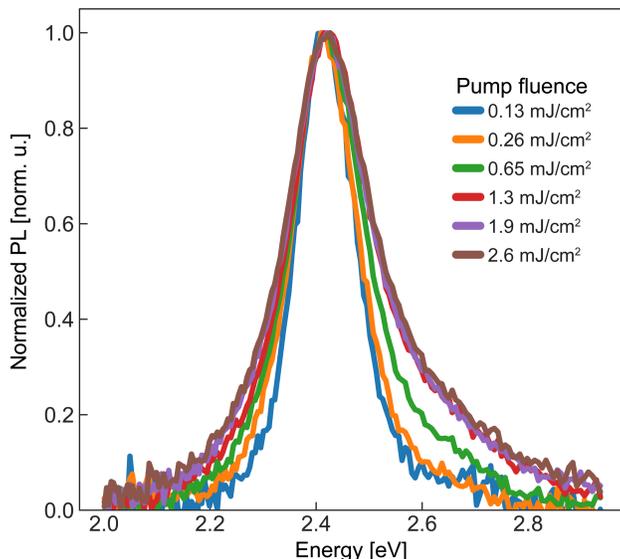


Figure S15. Integrated photoluminescence (normalized) for different pump fluences. Slight shift in peak position as well as broadening of the lineshape is observable. The broadening of the low-energy portion of the spectrum can be explained by formation of charged and multi-excitons in the nanocrystals, while the increase of the high energy side can be explained by emission from the states above the band edge.

Decay Rate Distribution

The calculated decay rate distribution is shown in Figure S16 for a full range of the energies and decay rates.

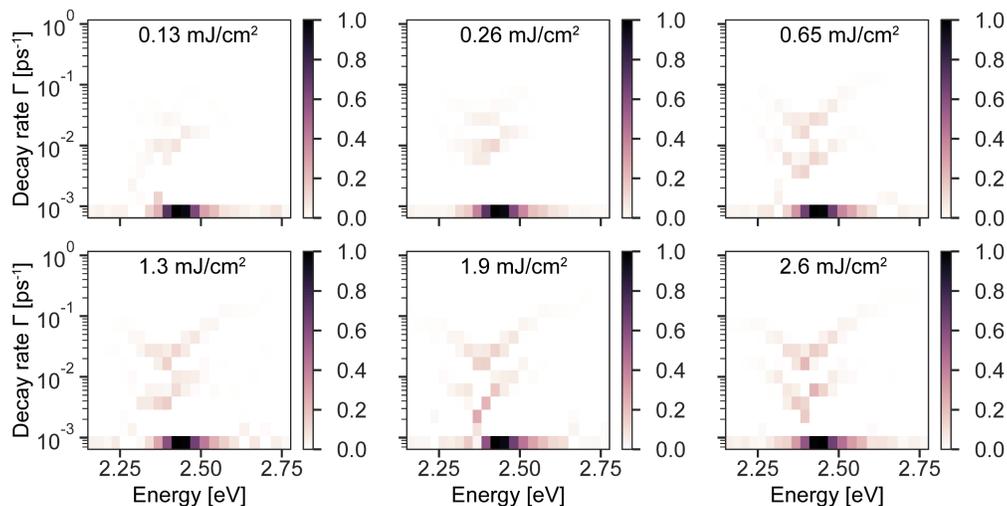


Figure S16. Calculated decay rate distribution for different pump fluences. The distribution is normalized as described in the text.

The Biexciton Binding Energy

The calculated decay rate distribution from Figure S16 can be used to study spectral properties of different luminescence species in CsPbBr₃ nanocrystals. By integrating over the decay rate axis, taken into account the appropriate limits for each of the X, X* and XX species, we can obtain the PL lineshapes of those species. Figure S17 shows the result of the integration over decay rate axis for the pump fluence of 2.6 mJ/cm².

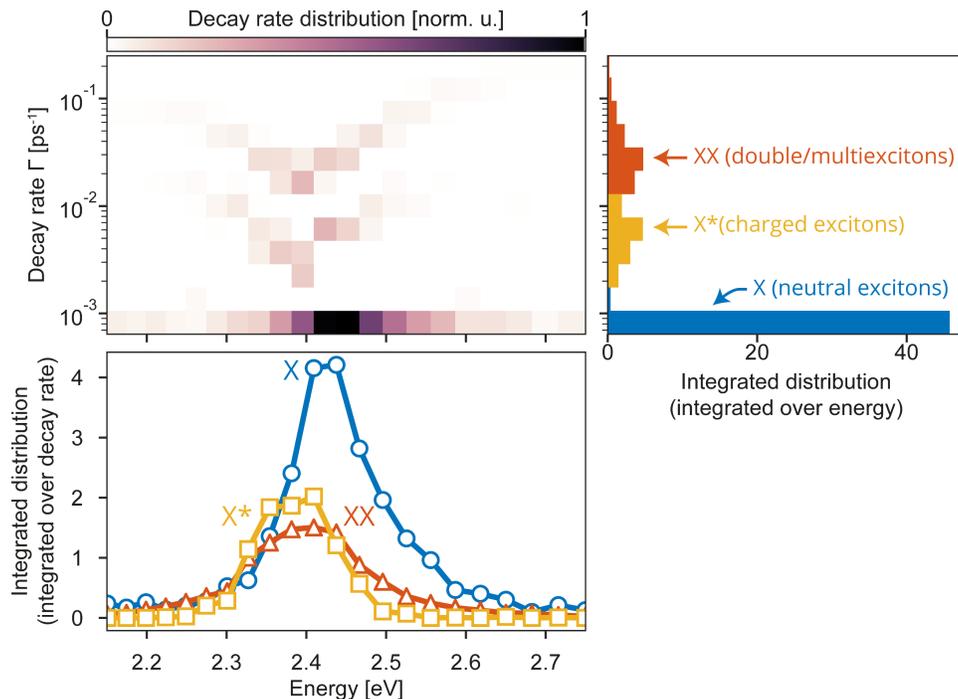


Figure S17. Detailed analysis of energy-resolved decay rate distribution obtained from FLUPS measurement at pump pulse energy of 100 nJ. By appropriate integration along decay rate axis, we can obtain spectral information about each individual luminescent species (bottom panel). This can be used to extract information about the biexciton binding energy. If we integrate over the energy axis, however, we can obtain decay rate distribution for each individual luminescent species in the system (right panel).

The single exciton lineshape is centered ~ 2.45 eV while XX peak has its maximum at ~ 2.4 eV, therefore yielding biexciton binding energy of ~ 50 meV. This value is in order of the previously measured biexciton binding energy of similar nanocrystals by Castaneda et al.¹³. By repeating this process for all the pump fluences measured, we can observe formation of charged excitons and multiexciton complexes, which is shown in Figure S18.

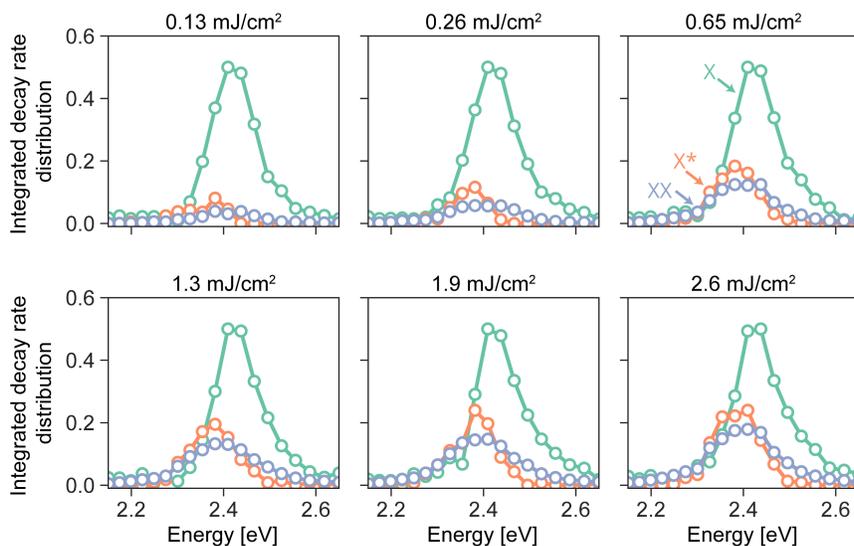


Figure S18. Photoluminescence lineshapes for three luminescent species identified in our study of CsPbBr₃ nanocrystals. Lineshapes are obtained by integrating energy-resolved decay rate distribution over the decay rate axis (please see Figure S17). In the regime of lowest fluence available (0.13 mJ/cm²), multiexcitons can be hardly resolved, while with increasing the fluence, its formation is readily resolved.

Integrating the energy resolved decay rate distribution over the energy axis, one can obtain distribution of decay rates for each of the luminescence species independently. This approach can be repeated for all the pump fluencies studied, which is shown in Figure S19.

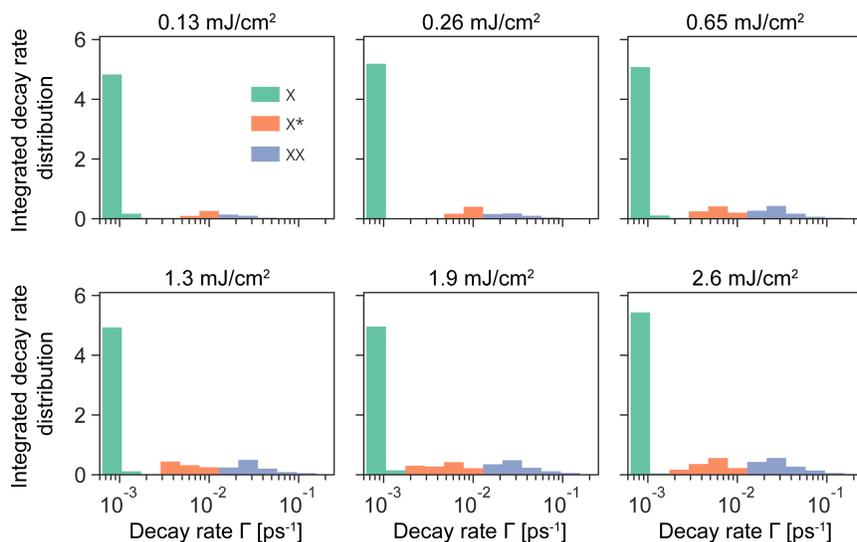


Figure S19. Decay rate distribution for three individual species identified in our study, obtained by integrating calculated energy-resolved decay rate distribution over the energy axis (see Figure S17).

Calculation Using Higher Resolution

We use an increased resolution to confirm that there are only three resolvable luminescent species in the two-dimensional map showing an energy-resolved decay rate distribution. Results obtained for all pump fluences doubling the resolution compared to Figure S16 is shown in Figure S20. There are no additional clusters observable in the high-resolution two-dimensional map. This result confirms the conclusion from results using lower resolution that there are only three luminescent species present.

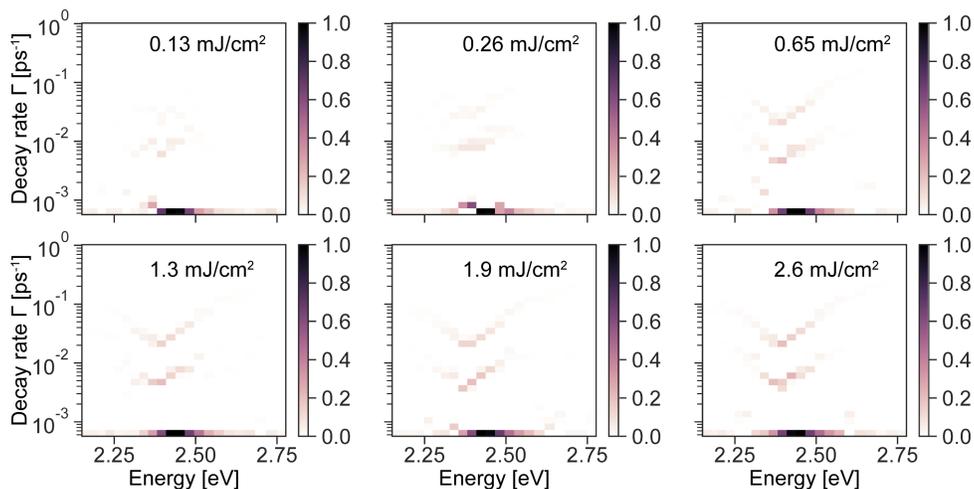


Figure S20. Calculated decay rate distribution for different pump fluences, using twice as high resolution compared to the results presented in Figure S16. The distribution is normalized as described in the text.

References

1. Santra, K.; Zhan, J.; Song, X.; Smith, E. A.; Vaswani, N.; Petrich, J. W., What Is the Best Method to Fit Time-Resolved Data? A Comparison of the Residual Minimization and the Maximum Likelihood Techniques As Applied to Experimental Time-Correlated, Single-Photon Counting Data. *The Journal of Physical Chemistry B* **2016**, *120* (9), 2484-2490.
2. Santra, K.; Smith, E. A.; Petrich, J. W.; Song, X., Photon Counting Data Analysis: Application of the Maximum Likelihood and Related Methods for the Determination of Lifetimes in Mixtures of Rose Bengal and Rhodamine B. *The Journal of Physical Chemistry A* **2017**, *121* (1), 122-132.
3. van Driel, A. F.; Nikolaev, I. S.; Vergeer, P.; Lodahl, P.; Vanmaekelbergh, D.; Vos, W. L., Statistical analysis of time-resolved emission from ensembles of semiconductor quantum dots: Interpretation of exponential decay models. *Physical Review B* **2007**, *75* (3).
4. Kapusta, P.; Wahl, M.; Erdmann, R., *Advanced Photon Counting*. 2015.
5. Berberan-Santos, M. N.; Bodunov, E. N.; Valeur, B., Mathematical functions for the analysis of luminescence decays with underlying distributions 1. Kohlrausch decay function (stretched exponential). *Chemical Physics* **2005**, *315* (1-2), 171-182.
6. Bertero, M.; Pike, E. R., Exponential-sampling method for Laplace and other dilationally invariant transforms: I. Singular-system analysis. *Inverse Problems* **1991**, *7* (1), 1.
7. Fletcher, G. C.; Ramsay, D. J., Photon-Correlation Spectroscopy of Polydisperse Samples: I. Histogram Method with Exponential Sampling. *Opt Acta* **1983**, *30* (8), 1183-1196.
8. Fletcher, G. C.; Ramsay, D. J., Photon Correlation Spectroscopy of Polydisperse Samples: II. Optimal choice of histogram parameters with exponential sampling. *Optica Acta: International Journal of Optics* **1986**, *33* (10), 1253-1262.
9. Ostrowsky, N.; Sornette, D.; Parker, P.; Pike, E. R., Exponential Sampling Method for Light Scattering Polydispersity Analysis. *Optica Acta: International Journal of Optics* **1981**, *28* (8), 1059-1070.
10. Protesescu, L.; Yakunin, S.; Bodnarchuk, M. I.; Krieg, F.; Caputo, R.; Hendon, C. H.; Yang, R. X.; Walsh, A.; Kovalenko, M. V., Nanocrystals of Cesium Lead Halide Perovskites (CsPbX₃, X = Cl, Br, and I): Novel Optoelectronic Materials Showing Bright Emission with Wide Color Gamut. *Nano Lett* **2015**, *15* (6), 3692-6.
11. Yarema, M.; Yarema, O.; Lin, W. M. M.; Volk, S.; Yazdani, N.; Bozyigit, D.; Wood, V., Upscaling Colloidal Nanocrystal Hot-Injection Syntheses via Reactor Underpressure. *Chemistry of Materials* **2017**, *29* (2), 796-803.
12. Heejae, C.; Il, J. S.; Jin, K. H.; Wonhee, C.; Eunji, S.; Dongho, K.; Weon-Kyu, K.; Jiwon, K., Composition-Dependent Hot Carrier Relaxation Dynamics in Cesium Lead Halide (CsPbX₃, X=Br and I) Perovskite Nanocrystals. *Angewandte Chemie* **2017**, *129* (15), 4224-4228.
13. Castañeda, J. A.; Nagamine, G.; Yassitepe, E.; Bonato, L. G.; Voznyy, O.; Hoogland, S.; Nogueira, A. F.; Sargent, E. H.; Cruz, C. H. B.; Padilha, L. A., Efficient Biexciton Interaction in Perovskite Quantum Dots Under Weak and Strong Confinement. *ACS Nano* **2016**, *10* (9), 8603-8609.