Electronic Supplementary Information: Optical Properties of Fully Inorganic, Spacer-Free Core/Gradient-Shell CdSe/CdZnS Nanocrystals at the Ensemble and Single-Nanocrystal Levels

Justine Baronnier[†], Benoit Mahler[†], Olivier Boisron[†],
Christophe Dujardin[†], Florian Kulzer[†,*] and Julien Houel[†*]

September 23, 2021

Contents

[†] Univ. Lyon, Université Claude Bernard Lyon 1
   CNRS UMR5306, Institut Lumière Matière
    69622 Villeurbanne, France
[*] To whom correspondence should be addressed: julien.houel@univ-lyon1.fr, florian.kulzer@univ-lyon1.fr

1 Characterization of OA-Capped Nanocrystals During Synthesis

We monitored shell growth during synthesis of the OA-capped NCs by measuring absorbance spectra at each step, which are shown in Figure S1; the change in slope below 500 nm is a signature of the growth of the gradient shell. This optical characterization was complemented by transmission electron microscopy, see Figure S2 for example images.
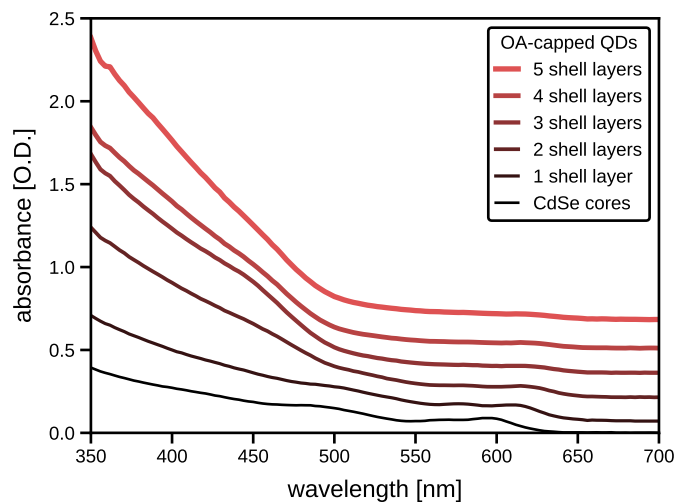
Figure S1: Absorbance spectra of OA-capped core/shell QDs at six stages of their synthesis, from the bare CdSe cores to the final core/gradient shell NCs (bottom to top, vertical offsets added for clarity). The absorbance spectrum of the CdSe cores was measured 180 s after injection of TOPSe, while the other five spectra were recorded at the end of the 10-minute annealing period for the respective shell layer.
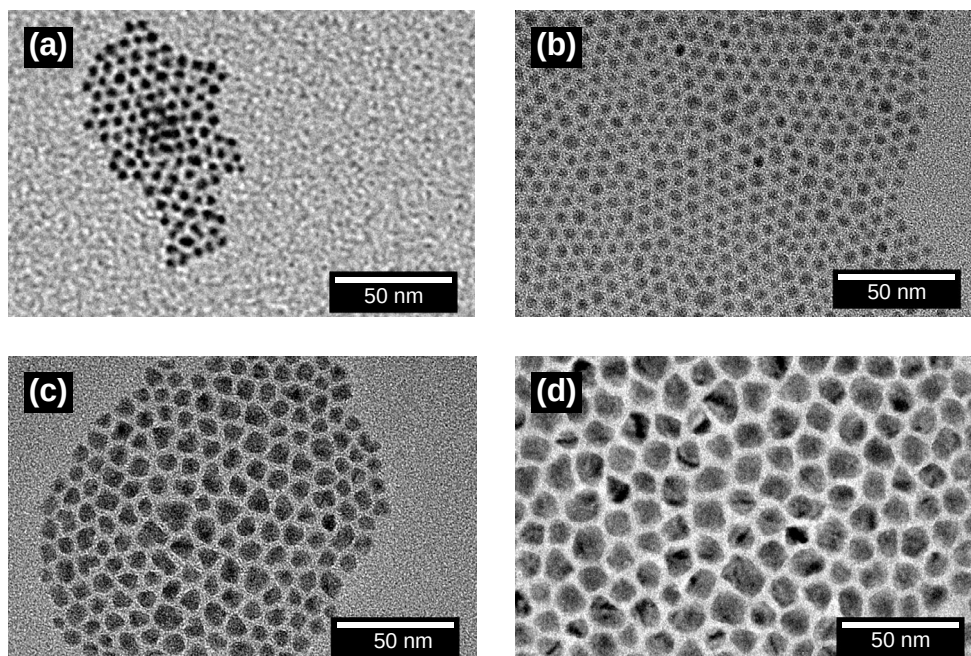
Figure S2: Transmission electron microscopy of OA-capped NCs at different stages of their synthesis: **(a)** bare CdSe cores; **(b)** after addition of the first shell layer; **(c)** after deposition of the second shell layer, which is richer in Zn; **(d)** final core/gradient shell NCs with 5 shell layers.

2 EDX and XPS Analysis of Halide-Capped Nanocrystals

Figure S3 (a) shows a scanning electron microscopy (SEM) image of a film of halide-capped NCs, prepared by letting dry a drop of NC solution on a titanium/gold-covered silicon wafer. Elemental analysis of this film and of a bare-substrate region was performed by energy-dispersive X-ray (EDX) spectroscopy at 15 keV (Phenom G2 Pro SEM). As can be seen in Fig. S3 (b), the NC film is, as expected, mainly composed of cadmium, zinc, sulfur and selenium, in addition to which we detected a clear signature of chlorine that confirms successful ligand exchange. It is furthermore important to note the absence of carbon and nitrogen, which helps to rule out that methyl ammonium chloride (CH_3NH_3Cl) was involved in the ligand exchange. The EDX spectrum of the reference region is dominated by silicon, gold and titanium, with weak contributions from scattered isolated NCs that are invisible in the SEM image of Fig. S3 (a).
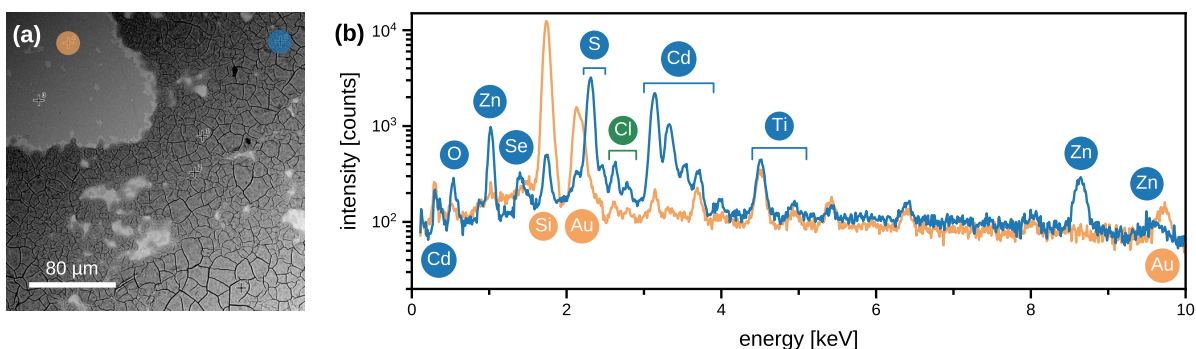
Figure S3: Elemental analysis of halide-caped NCs by energy-dispersive X-ray (EDX) spectroscopy. **(a)** SEM image of a film of halide-caped NCs on a titanium/gold-covered silicon wafer. **(b)** EDX spectra of the NC film (blue) and of the bare substrate (sandy brown); the two sample spots exhibiting these spectra are indicated in the SEM image (a) using the same color scheme. The elemental lines were annotated with the help of software provided for this purpose by the manufacturer of the SEM.

An X-ray photoelectron spectrum (XPS) of the halide-capped NCs was recorded with an Al source at an energy of 1.4866 keV (VG CLAM IV analyzer) and is presented in Fig. S4. Indexing of the spectral bands revealed the expected main elements (Cd, Zn, S, Se, Au) and a Cl 2p peak at 201 eV, which is this chlorine transition with the highest expected signal and confirms the presence of Cl on the NC surface; on the other hand, no sufficiently strong signal of bromine could be detected.
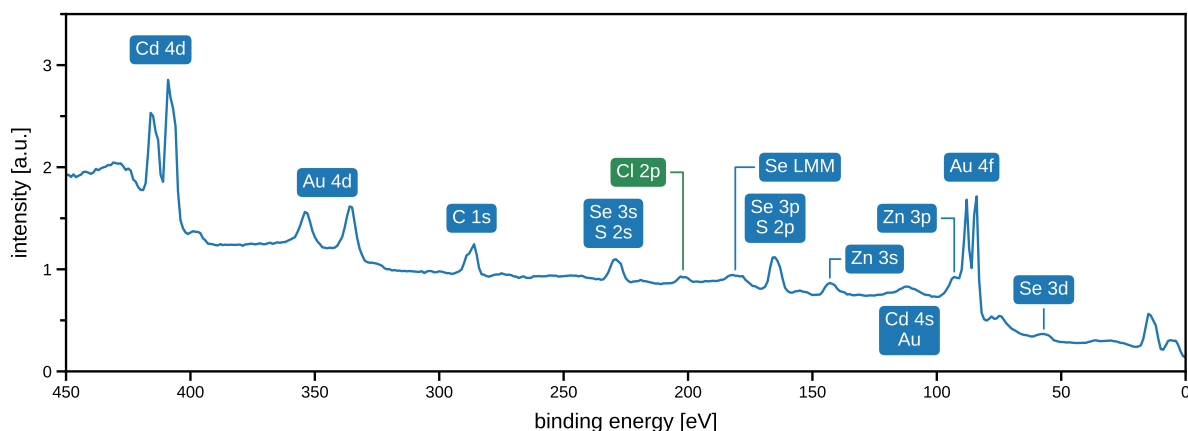
Figure S4: XPS spectrum of a film of halide-capped NCs on on a Ti/Au-covered Si wafer. The identified peaks are labeled by chemical element and electron orbital. (LMM denotes an Auger transition.)

3 Aging Phenomena in PL Spectra

The starting point for modeling our PL emission spectra is the band-gap energy E_{gap} of bulk CdSe, to which we add the energy of the lowest exciton state in a spherical quantum dot of radius R according to Brus, Eq. (12) in Ref. S11, to obtain the transition energy

$$E^*(R) = E_{\text{gap}} + \frac{\hbar^2 \pi^2}{2R^2} \left[\frac{1}{m_{\text{e}}^*} + \frac{1}{m_{\text{h}}^*} \right] - 1.786 \frac{e^2}{4\pi\epsilon_0 \epsilon R} \qquad , \tag{S1}$$

in which the second term is the sum of the ground-state energies of an electron and a hole with an effective mass of m_{e}^* and m_{h}^*, respectively, while the third term corresponds to the expectation value of the Coulomb attraction between these two charges in a medium with a dielectric constant of ϵ. As we are dealing with core/gradient-shell NCs, we interpret R in Eq. (S1) as an *effective* confinement radius that can be larger than r_{core}, the geometric radius of the NC core; we have furthermore omitted the solvation-energy term of Ref. S11 because neither the physical interpretation nor the mathematical expression of the underlying effect is applicable to a system without a well-defined interface between two distinct dielectrics. The numerical factor of the Coulomb term in Eq. (S1) arises from the integral

$$\begin{aligned}
\langle r_{12}^{-1} \rangle &= \frac{1}{4\pi^2 R^2} \int_0^R \int_0^R \int_{\Omega_1} \int_{\Omega_2} \sin^2(\pi r_1/R) \sin^2(\pi r_2/R) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \, d\Omega_2 \, d\Omega_1 \, dr_1 \, dr_2 \\
&= \frac{8}{R^2} \int_0^R \frac{\sin^2(\pi r_1/R)}{r_1} \int_0^{r_1} \sin^2(\pi r_2/R) \, dr_2 \, dr_1 \\
&= \frac{8}{\pi R} \int_0^\pi \frac{\sin^2(x_1)}{x_1} \int_0^{x_1} \sin^2(x_2) \, dx_2 \, dx_1 \\
&= \frac{1}{\pi R} \left[4 \int_0^\pi \sin^2(x_1) \, dx_1 - 2 \int_0^\pi \frac{\sin^2(x_1) \sin(2x_1)}{x_1} \, dx_1 \right] \\
&= \frac{1}{R} \left[2 + \frac{1}{2\pi} \int_0^\pi \frac{\sin(4x_1)}{x_1} \, dx_1 - \frac{1}{\pi} \int_0^\pi \frac{\sin(2x_1)}{x_1} \, dx_1 \right] \\
&= \frac{1}{R} \left[2 + \frac{1}{2\pi} \int_0^{4\pi} \frac{\sin(x)}{x} \, dx - \frac{1}{\pi} \int_0^{2\pi} \frac{\sin(x)}{x} \, dx \right] \\
&= \frac{1}{R} \left[2 + \frac{\text{Si}(4\pi)}{2\pi} - \frac{\text{Si}(2\pi)}{\pi} \right] \approx \frac{1}{R} \cdot 1.786 \qquad ,
\end{aligned} \tag{S2}$$

where we have omitted the details of rewriting $|\mathbf{r}_1 - \mathbf{r}_2|^{-1}$, the inverse of the electro-hole distance, by means of the generating function of the Lengendre polynomials $P_l(\cos\theta)$, where θ is the angle between $\mathbf{r}_1$ and $\mathbf{r}_2$; this method of separating the angular parts (Ω_1, Ω_2) of the above integral from its radial parts (r_1, r_2) is a standard technique for Coulomb integrals of hydrogen-like wave functions [S12].

The central exciton emission frequency $\nu^*(R)$ follows from Eq. (S1) as $\nu^*(R) = E^*(R)/h$, and we model the homogeneous emission spectrum of a single NC as a normalized Gaussian spectral probability density $\rho_{\text{hom}}(\nu \,|R)$ distributed around this frequency,

$$\rho_{\text{hom}}(\nu \,|R) = \frac{h}{\sqrt{2\pi} \, \sigma_0} \exp\left\{ -\frac{[\nu - \nu^*(R)]^2}{2 \, (\sigma_0/h)^2} \right\} \qquad , \tag{S3}$$

whose homogeneous linewidth σ_0, expressed here in units of energy difference, we expect to be on the order of the thermal energy, $\sigma_0 \approx k_{\text{B}} T$. To take into account the heterogeneity of

our QD samples, we assume that the effective confinement radius R of the individual NCs is governed by a normal probability density $p(R)$ centered on a certain average value R_0,

$$p(R) = \frac{1}{\sqrt{2\pi}\,\sigma_R}\,\exp\left[-\frac{(R-R_0)^2}{2\,\sigma_R^2}\right] \qquad . \tag{S4}$$

The width σ_R of this distribution could be influenced both by variations in NC size and by structural factors such as trapped charges or other defects modifying the confinement potential. We therefore expect σ_R of a newly-deposited sample to be comparable to the NC size distribution observed in TEM images (about $10\,\%$); structural degradation due to aging could then increase σ_R and/or shift R_0.

Combining the homogeneous emission spectrum of Eq. (S3) with the size distribution of Eq.(S4), we obtain the inhomogeneous PL emission spectrum $\rho(\nu)$ of an NC ensemble by numerical integration,

$$\rho(\nu) = \int_{r_{\mathrm{core}}}^{\infty} p(R)\,\rho_{\mathrm{hom}}(\nu\,|R)\,dR \qquad , \tag{S5}$$

where the lower integration boundary was set to $r_{\mathrm{core}} = 2.5\,\mathrm{nm}$, the core diameter of our QDs, because we assume that the NC core remains unaffected by structural modifications of the shell so that the exciton is never confined to a sphere smaller that the core. We have nevertheless verified that the choice of the lower integration boundary in Eq. (S5) does not significantly alter the predictions of our model, provided that this boundary remains large enough to avoid problems with the numerical integration algorithm arising from the divergence of Eq. (S1) at $R = 0$. We furthermore note that the Gaussian size distribution of Eq. (S4) does not stipulate any minium radius, which means that after fitting any given spectrum one has to make sure that the identified parameters R_0 and σ_R indeed imply only a negligible fraction of NCs with a confinement radius smaller than r_{core}; in practice, this requirement means having to verify that $\int_{-\infty}^{r_{\mathrm{core}}} p(R)\,dR \ll 1$.

The PL emission spectrum $S(\lambda)$ measured on the wavelength scale with a resolution of $\Delta\lambda$ follows from Eq. (S5) as

$$\begin{aligned} S(\lambda) &= \int_{c/(\lambda+\Delta\lambda/2)}^{c/(\lambda-\Delta\lambda/2)} \rho(\nu)\,d\nu \approx \rho(c/\lambda) \int_{c/(\lambda+\Delta\lambda/2)}^{c/(\lambda-\Delta\lambda/2)} d\nu \\ &= \rho(c/\lambda)\,\frac{c\,\Delta\lambda}{(\lambda+\Delta\lambda/2)(\lambda-\Delta\lambda/2)} \approx \rho(c/\lambda)\,\frac{c\,\Delta\lambda}{\lambda^2} \propto \frac{\rho(c/\lambda)}{\lambda^2} \qquad , \end{aligned} \tag{S6}$$

where we made use of the fact that $\Delta\lambda \ll \lambda$ and we have furthermore assumed that the emission spectrum varies slowly on the scale of $\Delta\lambda$, i.e.,

$$\left|\frac{d\rho(c/\lambda)}{d\lambda}\right| \cdot \Delta\lambda \ll \rho(c/\lambda) \qquad . \tag{S7}$$

The final proportionality relation of Eq. (S6) is sufficient for modeling the normalized PL emission spectra that we present in the main article; we simply applied the same normalization procedure to the simulated spectra.

To interpret our PL emission spectra, we used $E_{\mathrm{gap}} = 1.74\,\mathrm{eV}$, $m_{\mathrm{e}}^* = 0.13\,m_0$, $m_{\mathrm{h}}^* = 0.3\,m_0$, and $\epsilon = 10.2$ as fixed material constants [S13] in Eq. (S1), while the homogeneous linewidth σ_0, the average effective confinement radius R_0, and the width σ_R of the NC size distribution in Eqs. (S3) and (S4), respectively, were treated as adjustable model parameters. Our measured PL spectra show a weak background contribution decreasing linearly from $500\,\mathrm{nn}$ to $750\,\mathrm{nm}$, whose maximum amplitude is $0.02 - 0.05$ on the scale of the normalized spectra; we attribute this additional signal to the luminescence of residual by-products of the NC synthesis. We therefore corrected the PL spectra by subtracting a linear interpolation between the baselines
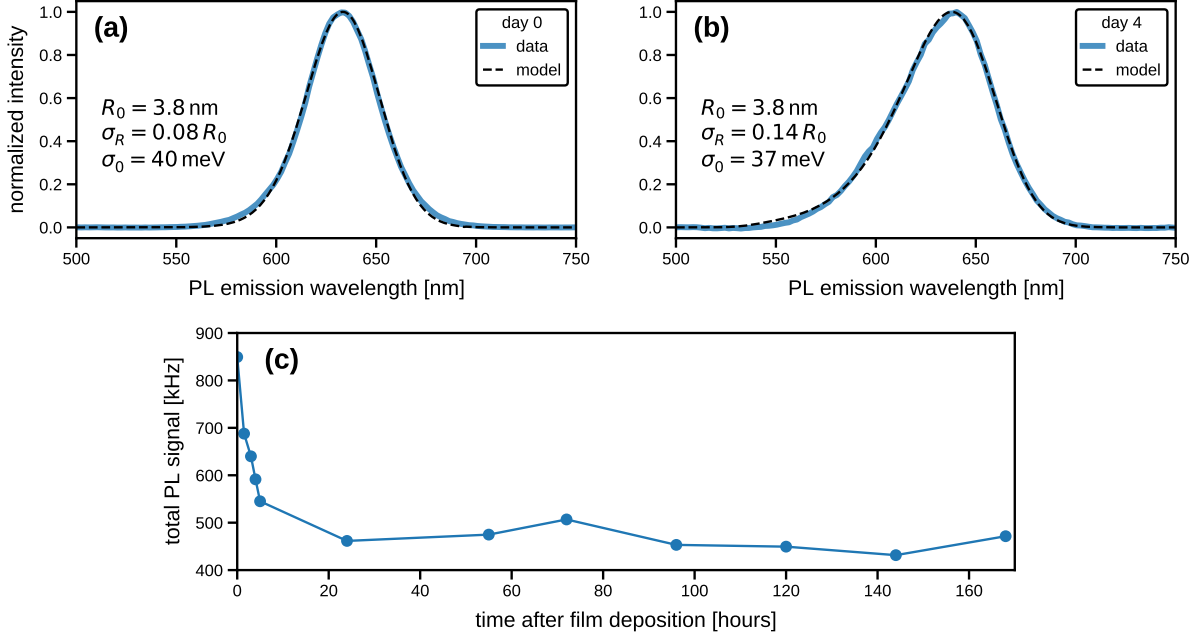
Figure S5: Photoluminescence emission spectra and total PL intensity of halide-capped NCs dispersed as a thin film on a glass substrate. The two PL spectra (solid lines) in the top row were recorded **(a)** on deposition day and **(b)** four days later; these spectra were derived from the ones in Fig. 2 c of the main article by subtraction of a weak linear background (see text for details) to facilitate comparison with the background-free model described in Eqs. (S1)–(S6). The dashed lines show least-squares-minimization adjustments to the spectra; the corresponding best-fit model parameters are indicated in each graph. **(c)** The total detected PL emission intensity as a function of time after film deposition. Filled circles indicate measurements, between which the film was kept in the dark but exposed to ambient air.

levels at $500\,\text{nm}$ and at $750\,\text{nm}$ (inferred by averaging 10 data points at each end) to obtain background-free spectra, which we then renormalized to a maximum intensity of unity. The resulting corrected spectra for the thin-film sample of halide-capped NCs are shown in Fig. S5, as are the model curves after least-squares optimization of the three adjustable parameters.

As can be seen in Fig. S5 a and b, Eqs. (S1)–(S6) reproduce the spectra of both the newly-deposited film and the four-days-old sample very well with an average effective confinement radius of $R_0 = 3.8\,\text{nm}$ and a homogeneous linewidth σ_0 around $40\,\text{meV}$ ($\approx 1.5\,k_\text{B}T$). In the framework of this model, the aging phenomenon is thus attributed to a widening of the distribution of the effective confinement radius, which is inferred to have a standard deviation of $8\,\%$ in the fresh film (in good agreement with the $10\,\%$ observed for the size distribution in TEM images) and is found to have increased to $14\,\%$ in the aged sample. We note that the imputed R_0 is, as expected, somewhat larger than the core radius ($2.5\,\text{nm}$) but stays well below the overall radius of the core/shell structure ($7.4\,\text{nm}$). Furthermore, the simplification of using a Gaussian size distribution, Eq. (S4), in a numerical integration starting at r_core, Eq. (S5), remains compatible with the results of the least-squares fits, as we are effectively only cutting off negligible fractions of the nominal distributions, about 10^{-5} for day 0 and $0.7\,\%$ for day 4, according the identified optimum values for R_0 and σ_R. Our simple model thus explains the observed aging phenomenon in terms of a moderate broadening of the distribution of effective confinement sizes; the increased asymmetry observed in the spectra of older samples is accounted for by the non-linear interplay of the R^{-1} and R^{-2} terms in the expression of the exciton energy, Eq. (S1). Fig. S5 c shows the evolution of the total emission intensity for identical excitation conditions; a rapid initial loss of about $50\,\%$ of PL intensity was observed, after which signal intensity and spectral shape remained stable for a week.

4 Maximum-Likelihood Analysis of Photoluminescence Decay Curves and Photon-Coincidence Histograms

We used maximum-likelihood (ML) analysis to identify a set of parameters leading to the best possible agreement between a given parameter-dependent model function and a measured set of data points. This data is made up of an ensemble $\{c_i\}$ of detected counts in n distinct detection channels, and it has to be compared to the corresponding predictions $\{g_i(\boldsymbol{\theta})\}$ of a model function g_i that depends on a vector of m adjustable parameters, $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)$. In our case the index $i = 1, \ldots, n$ of a detection channel represents the time that elapsed between the detection of the corresponding photon and a prior "start" event, which is an excitation laser pulse in fluorescence lifetime measurements and a previously-detected photon in photon-coincidence histograms. In the following sections, we will first present the general principle of maximum-likelihood analysis and then discuss its application to the various models that we used for our time-resolved photoluminescence data.

4.1 Quantifying Data-Model Agreement: Likelihood Function and Poisson Deviance

Maximum-likelihood (ML) analysis [S1–S5] is an efficient and unbiased approach to parameter estimation, which is preferable to the widely used χ^2 minimization, especially for counting channels containing a small number of photons [S6]. As outlined above, our data is comprised of a histogram of n channels, where each channel $i = 1, \ldots, n$ contains the number c_i of counts detected with a delay between $(i-1)\,t_{\mathrm{b}}$ and $i\,t_{\mathrm{b}}$ relative to the start event; t_{b} is the histogram bin time. A given model g_i has its associated vector $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)$ of m parameters such as decay rate(s) and amplitude(s). The likelihood function $\mathcal{L}$ is the joint probability of observing a given sequence of channel counts $\{c_i\}$ and can be written as [S3]

$$\mathcal{L}(c_1, \ldots, c_n | \boldsymbol{\theta}) = \prod_{i=1}^{n} p(c_i | \boldsymbol{\theta}) \qquad , \tag{S8}$$

where $p(c_i | \boldsymbol{\theta})$ is the conditional probability of detecting c_i counts in the i-th channel for a certain choice of model parameters $\boldsymbol{\theta}$. The specific parameter vector $\hat{\boldsymbol{\theta}}$ that maximizes the likelihood function $\mathcal{L}$ represents the maximum-likelihood estimation (MLE) of the model parameters for the data set $\{c_i\}$. For practical purposes it is usually $\ln(\mathcal{L})$ that is maximized by parameter variation, as the transition to logarithms allows to rewrite the product in Eq. (S8) as a sum of $\ln(p_i)$ terms. Analytic expressions for $\hat{\boldsymbol{\theta}}$ can be found in some cases by setting the derivatives of $\ln(\mathcal{L})$ with respect to the parameters θ_k to zero ($k = 1, \ldots, m$) and solving the resulting system of equations for the m unknown parameters; alternatively, a numerical search for a maximum in the m-dimensional $\boldsymbol{\theta}$ space can be performed if a given model is not amenable to analytic resolution of the ML conditions.

The multinomial probability of observing the data set $\{c_i\}$ for the n photon counting channels is given by [S5]

$$P(c_1, \ldots, c_n | \boldsymbol{\theta}) = \frac{N!}{\prod_{i=1}^{n} c_i!} \prod_{i=1}^{n} p_i(\boldsymbol{\theta})^{c_i} \qquad , \tag{S9}$$

where $N = \sum_{i=1}^{n} c_i$ is the total number of detected photons and $p_i(\boldsymbol{\theta})$ denotes the probability that any given photon will fall into detection channel i; these probabilities have to be normalized to the detection window such that $\sum_{i=1}^{n} p_i(\boldsymbol{\theta}) = 1$. The relationship between the model function $g_i(\boldsymbol{\theta})$ as defined above and the probabilities $p_i(\boldsymbol{\theta})$ in Eq. (S9) is straightforward, $g_i(\boldsymbol{\theta}) = N \cdot p_i(\boldsymbol{\theta})$. We used the multinomial model of Eq. (S9) because it offers certain advantages over an approach based on Poisson distributions [S5]. The probabilities $p_i(\boldsymbol{\theta})$ can be calculated for a given decay model as [S5]

$$p_i(\boldsymbol{\theta}) = \int_{\Delta_i} R(t | \boldsymbol{\theta})\, dt \qquad , \tag{S10}$$

where $\Delta_i = [(i-1)\,t_{\mathrm{b}},\, i\,t_{\mathrm{b}}]$ is the temporal interval associated with the i-th counting channel and $R(t|\boldsymbol{\theta})\,dt$ is the probability of detecting an emitted photon between t and $t + dt$ after a start event. [The specific models $R(t|\boldsymbol{\theta})$ and the resulting $p_i(\boldsymbol{\theta})$ for fitting our data are described in the following sections.] After calculating the probabilities p_i according to Eq. (S10), we used the fast-Fourier-transform-based routine `fftconv` of GNU Octave (www.octave.org) to convolute the model with the measured instrumental response function (IRF) to obtain the IRF-adjusted probabilities $\tilde{p}_i$ for our experimental setup; this convolution procedure does not preserve normalization, therefore we renormalized the probabilities once more to ensure $\sum_{i=1}^{n} \tilde{p}_i(\boldsymbol{\theta}) = 1$ before calculating $g_i(\boldsymbol{\theta}) = N \cdot \tilde{p}_i(\boldsymbol{\theta})$ and proceeding with the analysis. (The inclusion of the IRF in the models is an improvement compared to our earlier work [S7].) The single-detector IRFs were measured on attenuated reflections of the excitation laser pulses; the combined two-detector IRF needed for the photon-coincidence histograms was obtained by convolution of the IRFs of the two individual detectors.

Instead of maximizing the likelihood function $\mathcal{L}$ we used an equivalent approach that consists in finding $\hat{\boldsymbol{\theta}}$ by minimization of a statistical measure $D(\boldsymbol{\theta})$ that quantifies the discrepancy between the data set $\{c_i\}$ and the corresponding predictions $\{g_i(\boldsymbol{\theta})\}$ of the model. Given that the data in photon-counting histograms is Poisson-distributed, we used the Poisson deviance [S3]

$$D_{\mathrm{Poiss}}(\boldsymbol{\theta}) = 2 \sum_{i=1}^{n} \Big\{ c_i \, \ln\big[c_i/g_i(\boldsymbol{\theta})\big] - \big[c_i - g_i(\boldsymbol{\theta})\big] \Big\} \tag{S11}$$

as an appropriate statistical measure, which we minimized with the gradient-based `fminunc` search routine of GNU Octave to identify the MLE parameter vector $\hat{\boldsymbol{\theta}}$.

4.2 Models for Photon-Counting Data

4.2.1 Monoexponential Decay Curves

The probability density $R(t|\gamma, T)$ of a background-free monoexponential decay with rate γ is [S5]

$$R(t|\gamma, T) = \gamma \exp\big(-\gamma t\big) \, \frac{1}{1 - \exp(-\gamma T)} \qquad , \tag{S12}$$

which has been renormalized to the overall temporal width T of the detection window; in our case T is identical to the repetition time T_{rep} of the pulsed excitation laser. Given data acquisition with n channels of equal width (duration) $t_{\mathrm{b}} = T/n$, this means that the probability of a photon being detected in channel i, Eq. (S10), takes the following form [S5]:

$$p_i(\gamma, T, n) = \int_{(i-1)T/n}^{iT/n} R(t|\gamma, T)\, dt = \exp(-i\gamma T/n) \, \frac{\exp(\gamma T/n) - 1}{1 - \exp(-\gamma T)} \tag{S13}$$

In practice an additional adjustable parameter is needed, a delay time $\delta > 0$ to take into account the timing of the start event with respect to $t = 0$ of the measurement window. This delay shifts the integration boundaries in Eq. (S13), which now takes the form

$$p_i(\gamma, \delta, T, n) = \int_{(i-1)T/n-\delta}^{iT/n-\delta} \tilde{R}(t|\gamma, T)\, dt = \frac{\gamma}{1 - \exp(-\gamma T)} \int_{(i-1)T/n-\delta}^{iT/n-\delta} \exp\big[-\gamma \cdot \mathrm{mod}\,(t, T)\big] \qquad , \tag{S14}$$

where the modified probability density $\tilde{R}(t|\gamma, T)$ was obtained by replacing $t \to \mathrm{mod}\,(t, T)$ in Eq. (S12) in order to interpret negative times correctly, which correspond to the tail end of the decay curve triggered by the previous excitation pulse. (Recall that $T = T_{\mathrm{rep}}$ in our setup.) The expression of $p_i(\gamma, \delta, T, n)$ thus becomes slightly more complex than Eq. (S13) as one has

to distinguish the channels covering the time before, during, and after the transition from the previous to the current excitation pulse,

$$p_i(\gamma, \delta, T, n) = C \cdot \begin{cases} \left(e^{\gamma T/n} - 1\right) e^{\gamma[\delta - T(1+i/n)]} & \text{for } i \leqslant \lfloor n\delta/T \rfloor \\ \left[\left(e^{\gamma T(1/n-1)} - 1\right) e^{\gamma(\delta - Ti/n)} + 1 - e^{-\gamma T}\right] & \text{for } i = \lfloor n\delta/T \rfloor + 1 \\ \left(e^{\gamma T/n} - 1\right) e^{\gamma(\delta - Ti/n)} & \text{for } i \geqslant \lfloor n\delta/T \rfloor + 2 \end{cases} ,$$

(S15)

where $\lfloor \cdot \rfloor$ denotes the floor function and the normalization constant is $C = [1 - \exp(-\gamma T)]^{-1}$.

For monoexponential decay on non-zero background, the probability of photon detection in channel i changes to [S5]

$$p_i(\gamma, b, \delta, T, n) = \frac{b}{n} + (1 - b)\, p_i(\gamma, \delta, T, n) \quad , \tag{S16}$$

where b is the relative background contribution ($0 \leqslant b < 1$).

4.2.2 Bi-, Tri- and Quadri-Exponential Decay Curves

The probability of photon detection in the i-th channel for biexponential decay curves on non-zero background is [S5]

$$p_i(\gamma_1, \gamma_2, a, b, \delta, T, n) = \frac{b}{n} + (1 - b)\left[a\, p_i(\gamma_1, \delta, T, n) + (1 - a)\, p_i(\gamma_2, \delta, T, n)\right] \quad , \tag{S17}$$

where γ_1 and γ_2 are the two decay rates, a is the relative strength of the γ_1 component, and b the fraction of background counts.

The analogous expression for a triexponential decay is

$$p_i(\gamma_1, \gamma_2, \gamma_3, a_1, a_2, b, \delta, T, n) = \frac{b}{n} + (1 - b)\left[a_1\, p_i(\gamma_1, \delta, T, n) + a_2\, p_i(\gamma_2, \delta, T, n) \right.$$
$$\left. + (1 - a_1 - a_2)\, p_i(\gamma_3, \delta, T, n)\right] \quad , \tag{S18}$$

with the three decay rates γ_1, γ_2 and γ_3, the relative strengths a_1 and a_2 of the first two components, and the fraction of background counts b.

The quadri-exponential model takes the form

$$p_i(\gamma_1, \gamma_2, \gamma_3, \gamma_4, a_1, a_2, a_3, b, \delta, T, n) = \frac{b}{n} + (1 - b)\left[a_1\, p_i(\gamma_1, \delta, T, n) + a_2\, p_i(\gamma_2, \delta, T, n)\right.$$
$$\left. + a_3\, p_i(\gamma_3, \delta, T, n) + (1 - a_1 - a_2 - a_3)\, p_i(\gamma_4, \delta, T, n)\right] \quad , \tag{S19}$$

with the four decay rates γ_1, γ_2, γ_3 and γ_4, the relative strengths a_1, a_2 and a_3 of the first three components, and the fraction of background counts b.

4.2.3 Photon-Coincidence Histograms (Hanbury Brown – Twiss Measurement)

The types of nanocrystals investigated in this article are known to exhibit fluorescence antibunching in photon-coincidence measurements at the single-QD level [S8]. In an ideal background-free experiment on an individual QD, the probability of detecting a second photon (stop event) immediately after a first one has been registered (start event) is equal to zero, as each excitation-emission cycle requires a finite amount of time to complete, which leads to an antibunching dip with exponential flanks in the photon-coincidence histogram [S8]. For practical reasons a delay δ of roughly half the width of the detection window is imposed on

the stop channel (delayed-coincidence method), which leads to a mirror symmetry of the co-incidence histogram around $t = \delta$; the $t < \delta$ region of the antibunching curve corresponds to events for which the two detectors reversed their respective roles ("stop" actually happened before "start"). The corresponding model function is

$$R(t|\gamma, \delta, T) = \frac{\gamma}{e^{\gamma(\delta-T)} + e^{-\gamma\delta} + \gamma T - 2} \cdot \left[1 - \exp(-\gamma \cdot |t - \delta|)\right] \qquad , \qquad \text{(S20)}$$

which has been normalized to the duration T of the detection window. The probability for an observed start-stop delay to fall into the temporal interval covered by channel i is then given by

$$p_i(\gamma, \delta, T, n) = \int_{(i-1)T/n}^{iT/n} R(t|\gamma, \delta, T) \, dt \qquad . \qquad \text{(S21)}$$

Due to the absolute value $|t - \delta|$ in the formula for $R(t|\gamma, \delta, T)$ of Eq. (S20), the resulting expression for p_i has different forms for channels i before, at, and after the turning point (*i. e.*, the channel covering the moment $t = \delta$),

$$p_i(\gamma, \delta, T, n) = C \cdot \begin{cases} \left[\gamma T/n - \left(1 - e^{-\gamma T/n}\right) e^{\gamma(iT/n-\delta)}\right] & \text{for } i \leqslant \lfloor n\delta/T \rfloor \\ \left[\gamma T/n + e^{\gamma(\delta-iT/n)} + e^{\gamma[(i-1)T/n-\delta]} - 2\right] & \text{for } i = \lfloor n\delta/T \rfloor + 1 \\ \left[\gamma T/n + \left(1 - e^{\gamma T/n}\right) e^{\gamma(\delta-iT/n)}\right] & \text{for } i \geqslant \lfloor n\delta/T \rfloor + 2 \end{cases} \qquad , \quad \text{(S22)}$$

where the normalization constant is $C = [e^{\gamma(\delta-T)} + e^{-\gamma\delta} + \gamma T - 2]^{-1}$.

4.3 Uncertainty of Estimated Parameters: Bootstrap Method

We used the bootstrap method [S9] to estimate the uncertainties of the parameters that we obtained from the MLE analyses that we carried out. The bootstrap approach consists in creating synthetic data sets from one given measurement ensemble $\{c_i\}$ by randomly selecting n points from the experimental data. This selection process is implemented as drawing with replacement, which means that each data point has the same probability to be chosen in each draw, irrespective of whether it has already been selected in earlier draws. Each synthetic data set derived in this manner will therefore have a random fraction of the original data points missing ($\sim 37\%$ on average), while a number of the included data points is duplicated (and, on increasingly rare occasions, triplicated, quadrupled, etc.) such that the overall length n of the original data set is conserved. The ensemble of synthetic data sets thus obtained is then subjected to the same analysis as the measured data; the resulting distributions of the fit parameters furnish a good estimation for how precisely the parameters can be known from the data, provided that the measured data points can be assumed to be independent and identically distributed [S9]. The uncertainties of all MLE parameters reported in this article have been determined in this manner, based on 100 bootstrap runs (original data plus 99 resampled sets). These hundred ML-estimations of the model parameters $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)$ were used to calculate the covariance matrix of $\hat{\boldsymbol{\theta}}$, which yielded both the variances σ_k^2, $k = 1, \ldots, m$, of each individual ML-estimation $\hat{\theta}_k$ (and thus the standard deviations σ_k as estimates of the parameter uncertainties) and the covariances σ_{kl} between all pairs of parameters $\hat{\theta}_k$ and $\hat{\theta}_{l \neq k}$. The estimation of the uncertainty $\tilde{\sigma}_i$ of the ML-model at each data point i was then obtained by error propagation [S10] as

$$\tilde{\sigma}_i^2 = \sum_{k=1}^{m} \left(\frac{\partial g_i(\boldsymbol{\theta})}{\partial \theta_k}\bigg|_{\hat{\boldsymbol{\theta}}}\right)^2 \cdot \sigma_k^2 + 2 \sum_{k=1}^{m-1} \sum_{l=k+1}^{m} \left(\frac{\partial g_i(\boldsymbol{\theta})}{\partial \theta_k}\bigg|_{\hat{\boldsymbol{\theta}}}\right) \cdot \left(\frac{\partial g_i(\boldsymbol{\theta})}{\partial \theta_l}\bigg|_{\hat{\boldsymbol{\theta}}}\right) \cdot \sigma_{kl} \qquad , \qquad \text{(S23)}$$

where the partial derivatives of the model $g_i(\boldsymbol{\theta})$ with respect to the parameters θ_k were calculated numerically after convolution of the appropriate analytic expression of the model function with the measured instrumental response function.

4.4 Quantitative Comparison of Models: Likelihood-Ratio Test

If there is more than one candidate model to interpret a given dataset, maximum-likelihood analysis can be combined with a statistical test for the relative plausibility of the competing models. For nested models, *i. e.*, if one candidate is a special case of an alternative model, a likelihood-ratio test can be used to judge if the more complicated model achieves a statistically significant improvement in the description of the data [S3]. In the context of this work, this criterion was used to decide if the introduction of an additional decay component is justifiable; for example, a monoexponential decay, Eq. (S16), is a special case of a biexponential decay, Eq. (S17), with $a = 0$ or $a = 1$. Expressed in terms of Poisson deviances, Eq. (S11), a decision criterion can be formulated as follows [S3]: If a model with m adjustable parameters is a special case of an alternative model with $m + k$ parameters, then the more complex model will always fit the data better due to the additional k degrees of freedom providing more flexibility for adapting to the effects of experimental noise. If the simpler model is actually correct/sufficient, then the difference in the Poisson deviances of the two models is expected to follow a χ^2-distribution with k degrees of freedom,

$$(D_m - D_{m+k}) \sim \chi_k^2 (D_m - D_{m+k}) \qquad . \tag{S24}$$

One can therefore calculate a p-value for the statistical significance of an observed improvement in data-model agreement,

$$p = 1 - Q_k(D_m - D_{m+k}) \qquad , \tag{S25}$$

where $Q_k(x) = \int_0^x \chi_k^2(x') \, dx'$ is the cumulative distribution function associated with $\chi_k^2(x)$. After setting a threshold for statistical significance, for example $p = 0.05$, one can thus decide if choosing the more complex model is warranted. To give a practical example, if one wants to address the question of whether an additional decay component is necessary to describe a fluorescence lifetime curve, one has $k = 2$ as one would add two adjustable parameters, the supplementary decay constant and its relative contribution (amplitude). Choosing a threshold of significance of $p = 0.01$, one then finds that an improvement of at least $(D_m - D_{m+2}) = 9.2$ can be considered significant. Adopting this criterion would therefore amount to accepting an average of one false positive (choosing the model with an additional decay process where this is not warranted) in every 100 decisions taken.

5 Two-Level Saturation Model

Figure 4 of the main article shows the saturation effects for the number of detected photons as a function of excitation power; here we discuss a simple two-level saturation model for this phenomenon, illustrated in Fig. S6 for both resonant and non-resonant excitation. We consider the electronic ground state $|1\rangle$ and a given excited state $|2\rangle$, whose spontaneous relaxation rate (the inverse of its lifetime) is γ. The laser-induced pump rate W as a function of the excitation power P is given by

$$W = \frac{\sigma I}{h\nu} = \frac{\sigma P}{A h\nu} \qquad , \qquad (S26)$$

where σ is the cross-section of the $|1\rangle \leftrightarrow |2\rangle$ transition, ν is the average laser frequency, and $I = P/A$ is the laser intensity; A is an (effective) excitation area. Introducing the saturation power P_{sat} as

$$P_{\text{sat}} = \frac{kAh\nu}{\sigma} \qquad , \qquad (S27)$$

allows to rewrite the pump rate as $W = \gamma\,(P/P_{\text{sat}})$; the saturation power P_{sat} is the power at which the pump rate W equals the excited-state relaxation rate γ. In the case of resonant excitation, the laser induces transitions in both senses, $|1\rangle \to |2\rangle$ (absorption) and $|2\rangle \to |1\rangle$ (stimulated emission); the latter process is absent for non-resonant excitation.

The resonant excitation scheme of Fig. S6 translates into the following rate equations for the population densities N_1 (ground state) and N_2 (excited state):

$$\frac{dN_1}{dt} = -WN_1 + (\gamma + W)N_2 \qquad \text{and} \qquad \frac{dN_2}{dt} = +WN_1 - (\gamma + W)N_2 \qquad , \qquad (S28)$$

whose solution for the excited state population is

$$N_2(t) = N\,\frac{W}{2W + \gamma}\left\{1 - \exp\left[-(2W + \gamma)t\right]\right\} \qquad , \qquad (S29)$$

provided that the entire QD population started out in the ground state at $t = 0$, meaning that $N_1(0) = N$, where N is the total population density.

If we approximate the excitation laser as a rectangular pulse starting at time $t = 0$ and lasting until $t = T_{\text{pulse}}$, in which case the laser power P has to be understood as a mean effective power, then we obtain the following expression for the excited-state population density the end of each laser pulse:

$$N_2(T_{\text{pulse}}) = N\,\frac{P/P_{\text{sat}}}{1 + 2P/P_{\text{sat}}}\left\{1 - \exp\left[-(1 + 2P/P_{\text{sat}})\gamma\,T_{\text{pulse}}\right]\right\} \qquad , \qquad (S30)$$
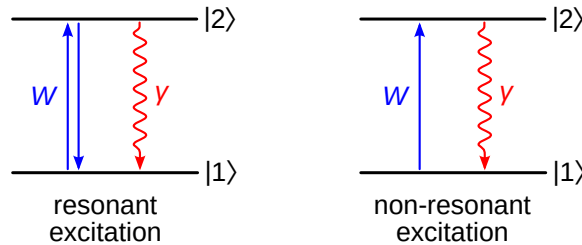
Figure S6: A two-level system as a simple model for the interaction of the excitation laser with a quantum dot, for resonant (left) and non-resonant(right) excitation. States $|1\rangle$ and $|2\rangle$ represent the ground and the excited state, respectively; W is the laser-induced pump rate and $\gamma = 1/\tau$ is the excited-state relaxation rate.

from which the expected number of detected photons, n_{phot}, can be obtained by multiplication with the detection volume V and the overall detection efficiency η. We thus deduce the power-dependent number of detected photons as

$$n_{\mathrm{phot}}(P) = \eta N V \, \frac{P/P_{\mathrm{sat}}}{1 + 2P/P_{\mathrm{sat}}} \left\{ 1 - \exp\left[-(1 + 2P/P_{\mathrm{sat}})\gamma \, T_{\mathrm{pulse}}\right] \right\} \quad , \qquad \text{(S31)}$$

which simplifies to a linear relationship in the low-power limit,

$$n_{\mathrm{phot}}(P) \approx \frac{\eta N V \left(1 - e^{-\gamma T_{\mathrm{pulse}}}\right)}{P_{\mathrm{sat}}} \, P \qquad\qquad \text{for } P \ll P_{\mathrm{sat}} \qquad . \qquad \text{(S32)}$$

If we repeat the same reasoning for the non-resonant excitation scheme, we obtain a slightly different expression for the number of detected photons,

$$n'_{\mathrm{phot}}(P) = \eta N V \, \frac{P/P_{\mathrm{sat}}}{1 + P/P_{\mathrm{sat}}} \left\{ 1 - \exp\left[-(1 + P/P_{\mathrm{sat}})\gamma \, T_{\mathrm{pulse}}\right] \right\} \quad , \qquad \text{(S33)}$$

which, however, is indistinguishable from the resonant case, Eq.(S31), when used as a fit function for the data of Fig. 4 without an independent means of accurately determining the cross-section σ as well as the experimental parameters η, N, V and A; furthermore, the linear low-power limit of Eq. (S33) is identical to Eq. (S32).

## 6 Model Comparison for Photoluminescence Decay Curves in Solution

The present section provides a graphical illustration of the maximum-likelihood analysis techniques described in detail in Section 4 and the comparison of the four different models, based on the photoluminescence decays of QDs in solution already presented in Fig. 3 a of the main article. Figures S7 and S8 show the corresponding time-resolved photoluminescence data again for the OA- and the halide-capped NCs, respectively; the improvement in the data-model agreement that is achieved by adding an additional decay component can be appreciated directly by comparing the corresponding graphs and the normalized residues R_i, which were calculated from the data points $\{c_i\}$ and a given maximum-likelihood model $g_i(\hat{\boldsymbol{\theta}})$ as

$$R_i = \frac{c_i - g_i(\hat{\boldsymbol{\theta}})}{\sqrt{g_i(\hat{\boldsymbol{\theta}})}} \qquad . \tag{S34}$$

If a model describes the data adequately then the only remaining data-model differences $c_i - g_i(\hat{\boldsymbol{\theta}})$ will arise from photon-counting noise, meaning that the c_i are expected to be
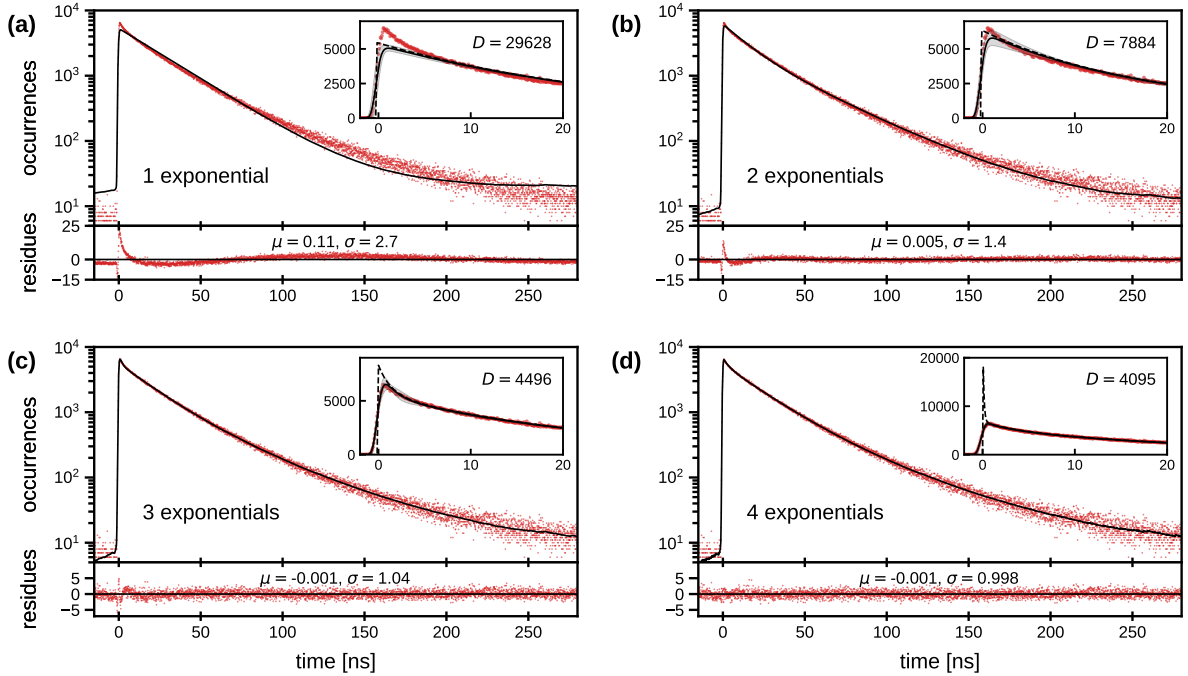
Figure S7: The photoluminescence decay curve of OA-capped NCs in solution (the same data is shown normalized in Fig. 3 a of the main article) fitted with a mono-, a bi-, a tri- and a quadri-exponential model, respectively. The four main panels depict the (identical) data (filled circles) with the best-fit model in question (solid curve); the insets show the same graphical material with a linear ordinate for the first 20 ns of the decay; additionally, the $10\,\sigma$ confidence interval of each model [gray area, see Eq. (S23)] and the shape of the model function before convolution with the response of the instrument (dashed line) are included. The D values given are Poisson deviances D_{Poiss} as defined in Eq. (S11). The bottom panels show the normalized residues of each model, Eq. (S34), with an indication of the mean μ and the empirical standard deviation σ. The maximum-likelihood model parameters are: **(a)** monoexponential model, Eq. (S16) with decay time $\tau = 1/\gamma = (27.5 \pm 0.2)\,\mathrm{ns}$ and background contribution $b = (3.0 \pm 0.1)\,\%$; **(b)** biexponential model, Eq. (S17) with $\tau_1 = 1/\gamma_1 = (37.3 \pm 0.7)\,\mathrm{ns}$, $a = (64.9 \pm 2.6)\,\%$, $\tau_2 = 1/\gamma_2 = (14.4 \pm 0.7)\,\mathrm{ns}$, and $b = (1.3 \pm 0.1)\,\%$; **(c)** tri-exponential model, Eq. (S18) with $\tau_1 = (46.0 \pm 0.8)\,\mathrm{ns}$, $a_1 = (38.5 \pm 1.6)\,\%$, $\tau_2 = (20.5 \pm 0.3)\,\mathrm{ns}$, $a_2 = (59.0 \pm 1.6)\,\%$, $\tau_3 = (1.3 \pm 0.2)\,\mathrm{ns}$, and $b = (0.9 \pm 0.1)\,\%$; **(d)** quadri-exponential model, Eq. (S19) with $\tau_1 = (49.6 \pm 0.8)\,\mathrm{ns}$, $a_1 = (31.1 \pm 1.1)\,\%$, $\tau_2 = (22.3 \pm 0.3)\,\mathrm{ns}$, $a_2 = (64.7 \pm 1.1)\,\%$, $\tau_3 = (3.5 \pm 0.2)\,\mathrm{ns}$, $a_3 = (3.1 \pm 0.2)\,\%$, $\tau_4 = (0.2 \pm 0.1)\,\mathrm{ns}$, and $b = (0.8 \pm 0.1)\,\%$
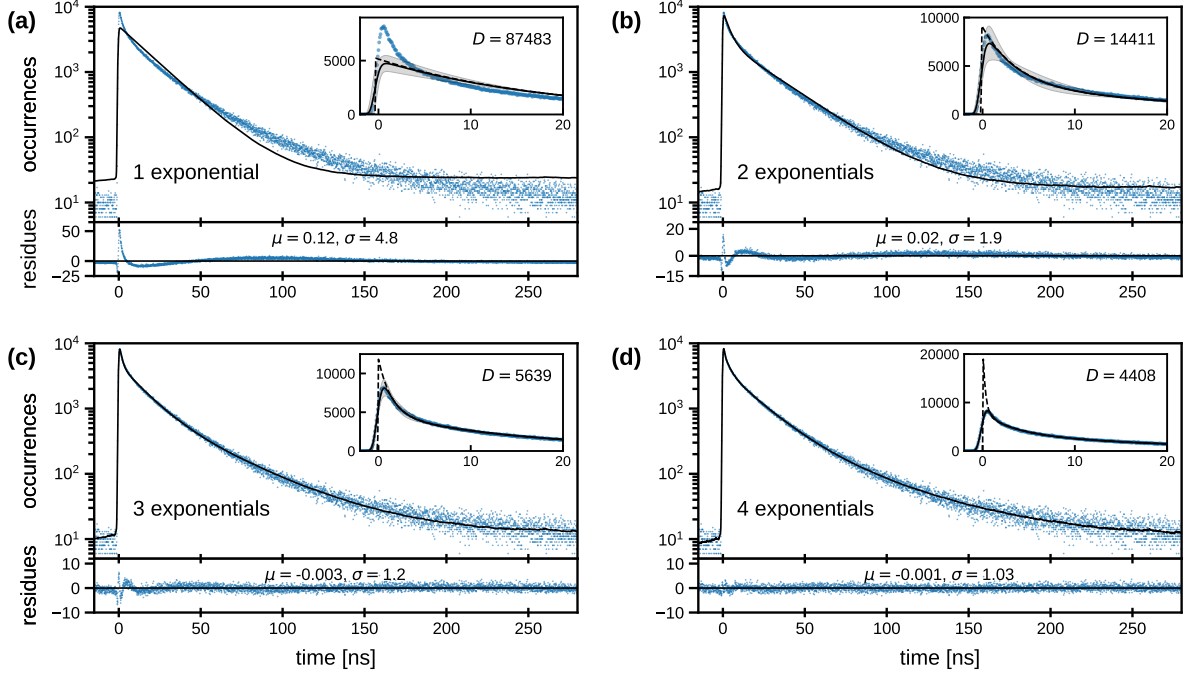
Figure S8: The photoluminescence decay curve of halide-capped NCs (the same data is shown normalized in Fig. 3 c of the main article) fitted with a mono-, a bi-, a tri-, and a quadri-exponential model, respectively. The four main panels depict the (identical) data (filled circles) with the best-fit model in question (solid curve); the insets show the same graphical material with a linear ordinate for the first 20 ns of the decay; additionally, the $10\,\sigma$ confidence interval of each model [gray area, see Eq. (S23)] and the shape of the model function before convolution with the response of the instrument (dashed line) are included. The D values given are Poisson deviances D_{Poiss} as defined in Eq. (S11). The bottom panels show the normalized residues of each model, Eq. (S34), with an indication of the mean μ and the empirical standard deviation σ. The maximum-likelihood model parameters are: **(a)** mono-exponential model, Eq. (S16) with decay time $\tau = 1/\gamma = (18.6 \pm 0.3)$ ns and background contribution $b = (6.1\pm0.1)\,\%$; **(b)** biexponential model, Eq. (S17) with $\tau_1 = 1/\gamma_1 = (25.3\pm0.3)$ ns, $a = (76.2\pm1.1)\,\%$, $\tau_2 = 1/\gamma_2 = (3.8 \pm 0.3)$ ns, and $b = (4.0 \pm 0.1)\,\%$; **(c)** triexponential model, Eq. (S18) with $\tau_1 = (37.4\pm0.9)$ ns, $a_1 = (39.2\pm1.5)\,\%$, $\tau_2 = (13.4\pm0.4)$ ns, $a_2 = (50.9\pm1.4)\,\%$, $\tau_3 = (1.4\pm0.1)$ ns, and $b = (2.7\pm0.1)\,\%$; **(d)** quadri-exponential model, Eq. (S19) with $\tau_1 = (47.3\pm1.0)$ ns, $a_1 = (23.8\pm0.8)\,\%$, $\tau_2 = (17.5\pm0.2)$ ns, $a_2 = (60.2\pm0.7)\,\%$, $\tau_3 = (3.3\pm0.2)$ ns, $a_3 = (12.4\pm0.3)\,\%$, $\tau_4 = (0.3\pm0.1)$ ns, and $b = (2.3 \pm 0.1)\,\%$

Poisson-distributed around $g_i(\hat{\boldsymbol{\theta}})$ with a standard deviation of $\sqrt{g_i(\hat{\boldsymbol{\theta}})}$ if the same measurement is repeated many times. Consequently, the normalized residues R_i calculated according to Eq. (S34) for all data points of any one measurement should have an average value of $\mu = 0$ and a standard deviation of $\sigma = 1$. Noticeably different values of μ and σ, as well as any systematic deviations from the zero-line in certain regions of the curve, indicate that the model in question fails to fully explain the data.

Figures S7 and S8 furthermore indicate the Poisson deviances D, Eq. (S11), for each model to provide an idea about how reduction of this statistical measure of data-model disagreement relates to the visible improvement of the description of the data; as mentioned in the main article, the statistical significance of the improvement achieved by adding one more decay component, Eq. (S25), is higher than 10^{-12} in all cases, as a difference in D of 56 suffices to surpass this significance threshold when comparing nested models with the likelihood-ratio test of Section 4.4.

7 Photon Antibunching in Single-QD Emission

As discussed in the main article, a decisive criterion for the observation of a single emitter is the detection of luminescence antibunching: The number of photon coincidences at lag time $\tau = 0$ is less than half of the limiting value at long times, which is an unambiguous signature of single-photon emission and proves that the signal does indeed originate from an individual emitter; Fig. S9 shows examples of this phenomenon for AO- and halide capped NCs. We systematically confirmed the presence of this antibunching dip (or the equivalent phenomenon under pulsed excitation) for the single-QD data that we present in the present work.
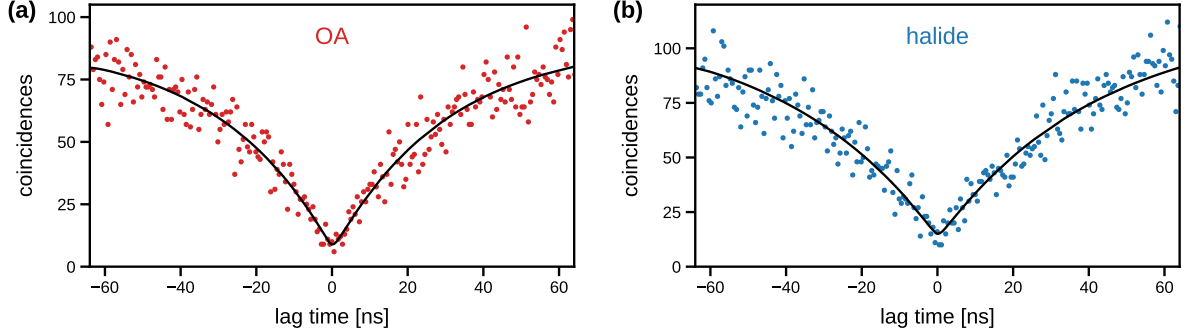
Figure S9: Antibunching in the photoluminescence emission of individual NCs: Photon coincidence histograms (filled dots) were measured with a two-detector Hanbury Brown – Twiss setup under continuous-wave excitation and fitted with the monoexponential model described in Section 4.2.3 (solid lines, including convolution with instrumental response). **(a)** OA-capped NC, $\tau_{\mathrm{rise}} = (30.2 \pm 2.4)\,\mathrm{ns}$; **(b)** halide-capped NC, $\tau_{\mathrm{rise}} = (42.3 \pm 4.2)\,\mathrm{ns}$.

References

[S1] Hall, P.; Selinger, B. Better Estimates of Exponential Decay Parameters. *J. Phys. Chem.* **1981**, *85*, 2941–2946.

[S2] Hall, P.; Selinger, B. Better Estimates of Multiexponential Decay Parameters. *Z. Phys. Chem. Neue Fol.* **1984**, *141*, 77–89.

[S3] Bajzer, Z.; Therneau, T. M.; Sharp, J. C.; Prendergast, F. G. Maximum-Likelihood Method for the Analysis of Time-Resolved Fluorescence Decay Curves. *Eur. Biophys. J. Biophys. Lett.* **1991**, *20*, 247–262.

[S4] Köllner, M. Best Estimates of Exponential Decay Parameters and the Design of Single-Photon-Counting Experiments. *Inst. Phys. Conf. Ser.* **1992**, *1992*, 151–154.

[S5] Köllner, M.; Wolfrum, J. How Many Photons are Necessary for Fluorescence-Lifetime Measurements? *Chem. Phys. Lett.* **1992**, *200*, 199–204.

[S6] Kim, J.; Seok, J. Statistical Properties of Amplitude and Decay Parameter Estimators for Fluorescence Lifetime Imaging. *Opt. Express* **2013**, *21*, 6061–6075.

[S7] Aubret, A.; Pillonnet, A.; Houel, J.; Dujardin, C.; Kulzer, F. CdSe/ZnS Quantum Dots as Sensors for the Local Refractive Index. *Nanoscale* **2016**, *8*, 2317–2325.

[S8] Lounis, B.; Bechtel, H. A.; Gerion, D.; Alivisatos, P.; Moerner, W. E. Photon Anti-bunching in Single CdSe/ZnS Quantum Dot Fluorescence. *Chem. Phys. Lett.* **2000**, *329*, 399–404.

[S9] Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes: The Art of Scientific Computing*, 3rd ed.; Cambridge University Press: Cambridge, 2007.

[S10] Ku, H. H. Notes on Use of Propagation of Error Formulas. *J. Res. Natl. Bur. Stand., Sect. C* **1966**, *70*, 263–273.

[S11] Brus, L. E. Electron-Electron And Electron-Hole Interactions In Small Semiconductor Crystallites - The Size Dependence Of The Lowest Excited Electronic State. *J. Chem. Phys.* **1984**, *80*, 4403–4409.

[S12] Bransden, B. H.; Joachain, C. J. *Physics of Atoms and Molecules*, 2nd ed.; Pearson: New York, 2003.

[S13] Madelung, O., Rössler, U., Schulz, M., Eds. *II-VI and I-VII Compounds; Semimagnetic Compounds*; Landolt-Börnstein - Group III Condensed Matter; Springer–Verlag: Berlin, 1999; Vol. 41B; p 580.