ACS NANO
Cite This: ACS Nano XXXX, XXX, XXX−XXX
www.acsnano.org

ARTICLE

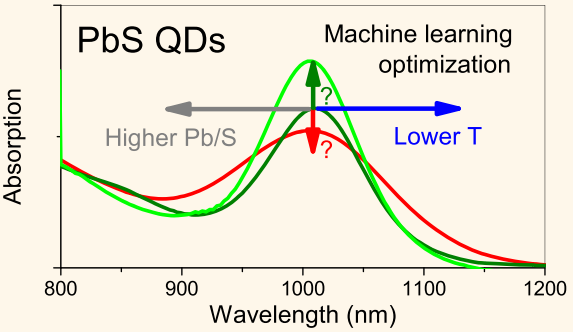
Machine Learning Accelerates Discovery of Optimal Colloidal Quantum Dot Synthesis

Oleksandr Voznyy,[†] Larissa Levina,[†] James Z. Fan,[†] Mikhail Askerka, Ankit Jain,[#] Min-Jae Choi, Olivier Ouellette, Petar Todorović, Laxmi K. Sagar, and Edward H. Sargent*

Department of Electrical and Computer Engineering, University of Toronto, Toronto, M5S 3G4, Canada

Supporting Information

ABSTRACT: **Colloidal quantum dots (CQDs) allow broad tuning of the bandgap across the visible and near-infrared spectral regions. Recent advances in applying CQDs in light sensing, photovoltaics, and light emission have heightened interest in achieving further synthetic improvements. In particular, improving monodispersity remains a key priority in order to improve solar cells' open-circuit voltage, decrease lasing thresholds, and improve photodetectors' noise-equivalent power. Here we utilize machine-learning-in-the-loop to learn from available experimental data, propose experimental parameters to try, and, ultimately, point to regions of synthetic parameter space that will enable record-monodispersity PbS quantum dots. The resultant studies reveal that adding a growth-slowing precursor (oleylamine) allows nucleation to prevail over growth, a strategy that enables record-large-bandgap (611 nm exciton) PbS nanoparticles with a well-defined excitonic absorption peak (half-width at half-maximum (hwhm) of 145 meV). At longer wavelengths, we also achieve improved monodispersity, with an hwhm of 55 meV at 950 nm and 24 meV at 1500 nm, compared to the best published to date values of 75 and 26 meV, respectively.**



KEYWORDS: *colloidal quantum dots, nanocrystals, synthesis, PbS, machine learning, Bayesian optimization*

PbS colloidal quantum dot (CQD) syntheses with polydispersities below 5−10% were established 15 years ago based on $PbOA_2$ and bis(trimethylsilyl) sulfide ($(TMS)_2S$) precursors.[1] Soon after, an alternative synthesis using $PbCl_2$ and elemental S dissolved in oleylamine (OLA) was developed.[2] Subsequent modifications allowed an extended range of sizes[3] and improved monodispersity[4] in the $PbCl_2$-based synthesis. Recently, a library of thiourea precursors has been developed for reproducible large-scale synthesis of PbS.[5]

Despite more than a decade of research, the synthetic parameter space has not been completely explored, and synthesis is usually guided by a simplified nucleation model.[6] Unfortunately, this model does not articulate the parameters that must be combined to achieve size focusing. It is known that increasing the Pb:S ratio, increasing the injection temperature, and increasing the OA:Pb ratio will increase the CQD size. However, up until now it was unclear which parameter combination is responsible for monodispersity and why smaller CQDs have consistently poorer size dispersion.

This limited understanding is related to complex behavior in the parameter space: retaining the bandgap requires changing two or more synthetic parameters at a time, and this complicates the exploration of the parameter space. Prior works have noted that an increased OA:Pb ratio, compensated by lower injection temperature, improves monodispersity;[7,8] however, using a high OA:Pb also limits the synthesis to larger CQD sizes (>1000 nm exciton peak).

We find herein that the addition of a growth-blocking agent, OLA, allows smaller nanoparticles to be synthesized that have a well-resolved exciton peak. Using machine learning (ML) methods, specifically Bayesian optimization implemented using a neural network, we build and explore a continuous numerical model of the parameter space to optimize the synthesis with monodispersity in mind. The model provides predictions regarding which parameter combination will achieve the desired properties, along with uncertainty associated with this prediction. Collecting additional experimental data following the predictions of the model allows for a gradual improvement in the accuracy of the model. Following this procedure, we find that OLA affects not only the size but also the monodispersity. Further improvements to monodispersity are achieved using a combination of high Pb:S ratio with lowered injection temperature, as well as by addition of metal chlorides.[9] We also discuss how the ML approach can provide insights beyond the original parameter space and contribute to generating strategies for further improvements.

Received: May 18, 2019
Accepted: September 20, 2019
Published: September 20, 2019

DOI: 10.1021/acsnano.9b03864
ACS Nano XXXX, XXX, XXX−XXX

RESULTS AND DISCUSSION

Synthesis of Smaller PbS CQDs. CQD synthesis is a competition among two factors: the nucleation rate of new particles and the growth rate of existing particles. Faster nucleation leads to consumption of the limited amount of precursors for new nuclei instead of growth, resulting in a larger number of smaller CQDs. If, however, the already-nucleated particles grow faster than the monomers can combine to form nuclei, fewer particles of larger size are obtained. It should be kept in mind that the nucleation does not necessarily stop completely, as assumed in the "burst nucleation" picture,[6] and may proceed even at the later stages of the reaction, resulting in a tail of small CQD sizes in the particle distribution.[10]

Excess OA in $PbOA_2$-based synthesis solubilizes nuclei, thus slowing nucleation and producing larger particles.[1] Thus, in order to achieve the smallest particles, we focus here on a synthesis that uses the minimum allowed OA:Pb ratio of ~2.4:1, just enough to solubilize PbO and convert it into $PbOA_2$ with only a small excess of OA.

To reduce the size further, we sought to speed nucleation or slow growth. Previously, Hens *et al.* demonstrated that the $PbCl_2$+S-based synthesis can be modified by the addition of trioctylphosphine sulfide (TOPS) as a low-reactivity sulfur source to obtain smaller particles.[3] This was rationalized by the reduced particle growth speed, whereas the nucleation was not affected due to the presence of a small amount of another more reactive precursor, OLA-S.

We sought to find a similar growth inhibitor for the $PbOA_2$+$(TMS)_2S$-based synthesis. In particular, the OLA used to solubilize S in the $PbCl_2$+S synthesis is known to become the primary ligand on the PbS surface, suggesting that it binds to Pb on the surface, though not more strongly than does OA.[3] We thus hypothesized that the presence of OLA can bind to, and block, the CQD surface, and in this way slow growth. It can, however, bind to $PbOA_2$ precursors, too, and thus slow nucleation. The binding strength may be sensitive to the nuclei/QD size, and therefore we sought to explore experimentally whether nucleation or growth was affected more. Our experimental tests demonstrated that smaller particles are achieved (down to 630 nm exciton peak), with the exciton peak well resolved in the absorption spectra, indicating that the slowing of particle growth was the larger effect.

Machine Learning Synthesis Optimization. We sought to analyze whether the synthesis could be further improved in terms of the exciton line width. We digitized laboratory data from the past 6 years (2300 syntheses shown in Figure 1, many of which were repeats of nominally the same parameter set). As can be seen, the parameter space remains largely unexplored because better monodispersity was not a primary goal until recently.

Optimizing synthesis is a laborious task because to stay at the same bandgap, not one but several synthetic parameters have to be changed at once. The local gradient along which to move in order to improve the desired output (in our case, the line width) can in principle be found by a finite difference method, *i.e.*, by doing multiple measurements with a small deviation around each point in the parameter space.

Estimating the local gradient is, however, complicated by experimental noise. Repeating the same synthesis several times results in bandgap variations of up to 30 nm (SI Figure 1),

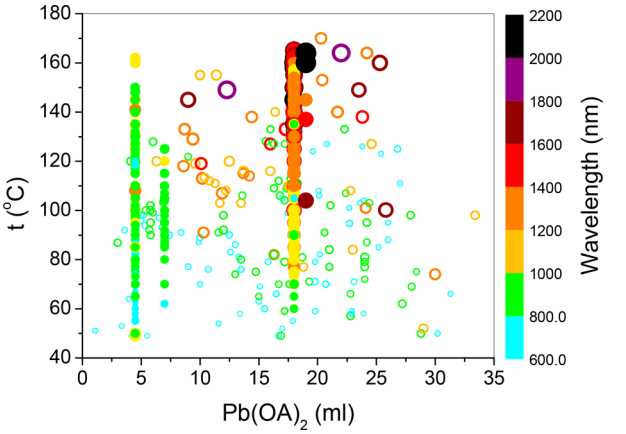


Figure 1. Parameter space coverage. Experimental data points in the input space of Pb precursor volume and injection temperature. Filled circles are the original pre-ML data (~2300 points); empty circles are the experiments performed following the ML suggestions (230 data points). Note that all data points with different S and OLA amounts are separated in other dimensions but are projected onto one plane in this plot. Color scale indicates the bandgap wavelength in nanometers.

rendering the estimate of the gradient unreliable. This variability originates from changes in the environment temperature and differences in glassware used for each experiment (microcracks, scuffs, microbubbles, variations in wall thickness, *etc.*, for nominally identical flasks), affecting the heating and cooling speed, the temperature of the solvent, and the temperature of the injected $(TMS)_2S$.

Testing various parameter combinations on a coarser grid provides an approach that is less susceptible to local noise and gives a broader picture of the effect of each parameter.[7] However, in light of the number of parameters and the density of the grid, exploring the entire parameter space may quickly become too time-consuming (SI Figure 2).

We turned to ML to replace the dense regular grid with a sparser set of points in the parameter space.[11] The results between the measured points are interpolated using smooth spline-like functions. In our case this is achieved using a neural network with nonlinear exponent-based activation functions. The accuracy of the interpolation is higher in proximity to available points, while confidence decreases when far from them. The general trends of the functional behavior are captured, and this enables suggesting new points in the parameter space that are likely to optimize the required metric of interest. The confidence interval for the model predictions can also be estimated (see Methods and SI Figure 1), allowing for a more robust exploration of the parameter space. Parameters suggested by the model are tested experimentally, and the results are added to the data set, allowing iterative improvement of the accuracy of the model in the regions where previously no data were available and confidence was low (SI Figure 1b–e and SI Figure 3), in a process of Bayesian optimization. Once the approximate position of the optimum is located, the algorithm starts to sample points around it on a denser and denser grid, until the exact optimum position is pinpointed. This nonlinear regression in multidimensional space has been proven to be at least an order of magnitude faster than the regular grid search[12] and has become more accessible with the advent of open source ML libraries (TensorFlow, SciKit, PyTorch, *etc.*). The potential of nonlinear

DOI: 10.1021/acsnano.9b03864
ACS Nano XXXX, XXX, XXX–XXX

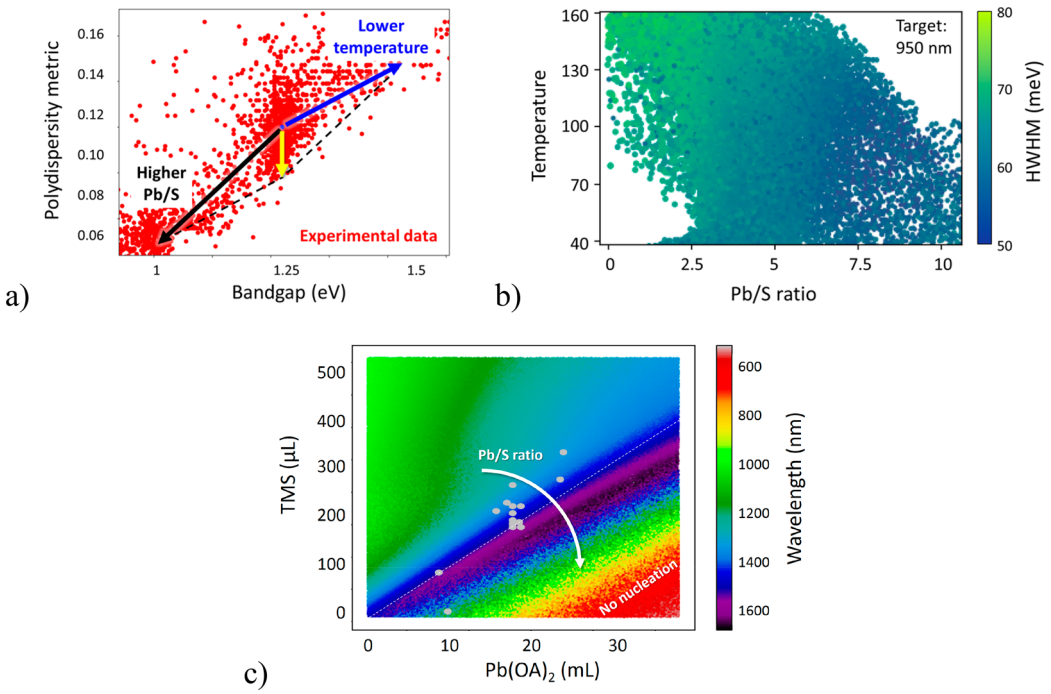


Figure 2. Factors affecting monodispersity. (a) Combined effect of increased Pb precursor amount and decreased injection temperature, allowing to maintain the bandgap but improve the monodispersity. See Methods for the definition of polydispersity metric. (b) Same as (a) but in the input space (darker color means smaller hwhm). (c) ML model predictions regarding the effect of the Pb:S precursor ratio on the bandgap, with temperature range limited to values optimal for growth of CQDs with 1500 nm exciton peak. Gray circles: experimental data points falling within the same parameter range.

regression in quantum dot synthesis has recently been demonstrated in selecting multiple synthetic procedures to target the same bandgap in perovskite CQDs.[13] We emphasize that the approach does not require thousands of data points, and several tens of syntheses may suffice to achieve meaningful improvements targeting a specific metric.

Factors Affecting Monodispersity. The trainedML model allowed us to visualize the effect of each synthetic parameter on the outcome of the synthesis, starting from a given combination of parameters (Figure 2a, SI Figure 4). This analysis shows that the effects of Pb:S ratio and of temperature are not equivalent—they are not parallel in the plot—and that the combination of higher Pb:S ratio and lower injection temperature is beneficial for half-width at half-maximum (hwhm) (Figure 2a yellow arrow and Figure 2b). This is similar to observations for $PbCl_2$+S-based syntheses in which Pb:S ratios as high as 24:1 were used to yield the best monodispersities. However, the effect of injection temperature on monodispersity was not previously explored, and instead the CQD size was controlled using an (operator-dependent) quenching of the reaction at a specific time.[2−4]

We also find that in the $PbOA_2$-based synthesis, the Pb:S ratio cannot be increased indefinitely, achieving optimum monodispersity at Pb:S ratios around 8:1, beyond which the nucleation essentially stops (red region in Figure 2c), resulting in a nearly transparent solution with a bandgap of ∼650 nm and poor monodispersity with no exciton peak discernible.

We plot in Figure 2c a continuous function of outputs in the parameter space, interpolating between experimental points. We observed that the reaction is sensitive not just to Pb:S ratio but also to the absolute concentration of the precursors, *i.e.*, saturation of the solution. This effect is even more pronounced in the synthesis of smaller CQDs.

Dependence on the concentration is a signature that the reaction remains in part diffusion-limited. This defies the assumptions inherent in an idealized model of nucleation and growth[6] and highlights the importance of systematic exploration of the parameter space instead of relying on simplified models.

Effect of OLA. Our initial efforts focused on using OLA exclusively for the smallest PbS CQDs. Only several exploratory syntheses were made that resulted in larger CQDs. These few points, however, were sufficient for the ML algorithm to detect a positive effect of OLA on the monodispersity and prompted us to explore this effect further in combination with other parameters. The effect of OLA is similar to that of reduced temperature (Figure 2a) and provides an additional degree of freedom in tuning the synthesis and allowing to further improve the monodispersity for CQDs in the 600−1300 nm wavelength range (Figure 3).

Effect of Chlorides. Addition of chlorides ($CdCl_2$, $PbCl_2$, tetrabutylammonium chloride) was previously used to improve surface passivation,[9,14,15] and we find here that it has an effect on monodispersity as well (SI Figure 5). The reason for such improvement is rationalized by an analysis of the trends using the ML model. Throughout our experiments we found that $PbCl_2$ has the strongest effect on monodispersity compared to other chlorides. In addition, in order to improve the solubility of metal halides, an addition of OLA was always used. The effect of chlorides is thus a combination of increasing the metal content and adding OLA, a combined effect like that of increased $PbOA_2$ and reduced injection temperature discussed above (Figure 2a). We note, however, that $PbCl_2$+OLA injection can further improve even the best synthesis where OLA addition has already been employed and its amount optimized. This means that the timing of Pb and OLA addition

DOI: 10.1021/acsnano.9b03864
ACS Nano XXXX, XXX, XXX−XXX

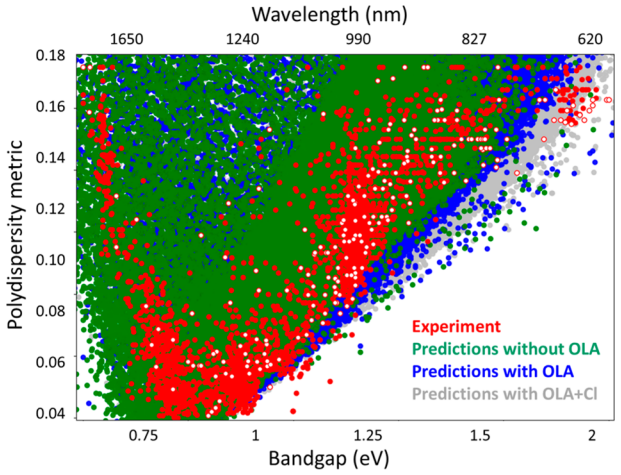


Figure 3. Effect of OLA and chlorides on synthesis. Experimental data (red points) and ML predictions without additives (green), with OLA (blue), and OLA plus Cl (gray) in the output space. Solid red circles denote pre-ML historical data, while open circles are the data points obtained in the course of ML optimization. Gray points encompass the widest range of possible outcomes and thus are placed on the lowest layer, followed by the blue and green. Experimental results (red) are placed on the topmost layer, overlapping with the predictions of the ML model. See Methods for the definition of the polydispersity metric and for an explanation of how this figure was generated.

plays a role, presumably by decoupling their effect on nucleation, which happens at early times, *vs* particle growth, which happens at later times. This indicates an additional synthetic parameter that can be optimized in the future.

Experimentally Verified Best Syntheses. Below we present the best syntheses for several CQD sizes in terms of monodispersity that were suggested by the ML model and verified experimentally. Each was performed multiple times, and error bars are reported along with the results.

Addition of OLA allows reproducibly achieving CQDs with a bandgap in the 610−620 nm range (~2 eV) with a well-defined excitonic peak (hwhm of 145 meV). Additional XRD and TEM characterization of these CQDs is provided in SI Figures 6 and 7, yet absorption remains the most sensitive technique to assess monodispersity.

The results for 950 nm (1.3 eV) are the best reported to date, with the best hwhm of the exciton peak of 55 meV closely approaching the estimated ~50 meV hwhm single-dot (homogeneous) emission line width.[16]

For CQDs around 1500 nm (0.83 eV), our hwhm of 24 meV is only 1−2 meV, better than the best previously reported using $PbCl_2$+S-based synthesis.[4] Nevertheless, we still observe a dramatic improvement in the peak-to-valley ratio (SI Figure 8), indicating that here too we are approaching the homogeneous line width. We also observe three well-defined higher-energy transition peaks in the absorption spectra (Figure 4) in contrast to only one peak observed for the best $PbCl_2$+S-based synthesis,[4] although this difference could be due to other factors, such as the CQD shape (SI Figure 7). Such CQDs are the best candidates for achieving dense packing and band-like electronic transport in CQD ensembles.[4,17] The true homogeneous line width, however, depends on multiple factors, such as electron−phonon-assisted broadening[18] and fine structure of multiple optical transitions contributing to the first exciton peak.[19] In the absence of single-dot line width measurements, it is thus difficult to assess whether we have achieved a homogeneous line width.

CONCLUSIONS

We have demonstrated that OLA allows improved $PbOA_2$+$(TMS)_2S$-based PbS CQD synthesis, in both size range and monodispersity.

Machine learning allowed more efficient exploration of the parameter space, a systematic route to finding optimal synthetic conditions and a method of analysis that led to suggestions of avenues toward further synthetic improvement. In particular, even higher Pb:S ratios, but with the use of continuous injection, could be explored to avoid nucleation blocking. Late injection of OLA could be attempted to decouple its effect on nucleation and growth. Different ligands could be used to decouple precursor solubility from viscosity, and the effect of OA:Pb ratio on monodispersity can be further explored.

We note that a similar optimization could be used to target other parameters of interest, as well as applied to other QDs or, more generally, to design-of-experiments optimization. For example, for a scale-up synthesis employing an external chiller with silicone oil to control the solution temperature, it is desirable to develop a synthesis with injection temperature not exceeding 140 °C to avoid the use of an expensive higher-grade synthetic oil. An ML-based model can suggest syntheses with such limitations in mind. Such an optimization is expected to be beneficial also to the recently developed PbS CQD synthesis employing thioureas as a precursor.[5]

More accurate control of the synthetic parameters (environment temperature, flask quality, programmed heating and cooling speed) can improve reproducibility and pinpoint the parameters that cause outlier syntheses to work out better than the average. Postsynthetic size focusing in the presence of excess OA can also be explored.[1,20] Addition of OLA or metal chlorides, dropwise addition of precursors, and effect of quenching time and cooling rate could all be further explored with the aid of ML.

Table 1. Experimentally Verified Synthetic Parameters for Different Bandgaps, Targeting the Best Monodispersity

target wavelength (nm)	$PbOA_2$ (mL)	ODE_{Pb} (mL)	OLA (mL)	T_{inj} (°C)	$(TMS)_2S$ (μL)	$ODE_{(TMS)_2S}$ (mL)	$PbCl_2$ (mM)	peak/valley	hwhm (meV)
620	3.35	15	0.54	53	168	9	0	1.27 ± 0.1	145 ± 4
780	13.5	15	0.23	76	220	16.3	0.3	2.1 ± 0.1	100 ± 3
950	17.4	15	0.4	109	206	9.6	0.3	3.8 ± 0.5	60 ± 5
1150	17.7	15	0.15	110	210	8	0.3	6.6 ± 0.4	42 ± 2
1330	18	15	0	125	240	9	0.3	8.5 ± 1	27 ± 2
1500	18	15	0	140	210	8	0.3	9 ± 1	25 ± 1

DOI: 10.1021/acsnano.9b03864
ACS Nano XXXX, XXX, XXX−XXX

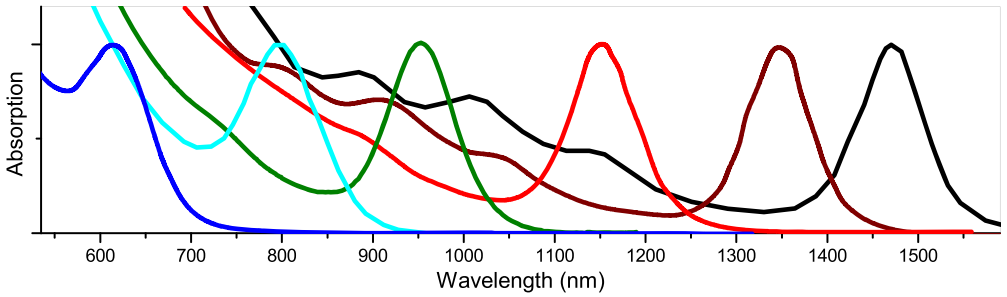


Figure 4. Absorption spectra for several CQD sizes with best monodispersities.

METHODS

$Pb(OA)_2$ Preparation. In a 250 mL round-bottom Schlenk flask, 9.0 g of PbO (99.99%, Alfa Aesar), 30 mL of oleic acid (90%, Caledon), corresponding to a 2.35:1 OA:Pb ratio, and 60 mL of 1-octadecene (95%, Caledon) was mixed and stirred at 110 °C overnight (over 10 h) under a vacuum of 2×10^{-2} Torr.

$(TMS)_2S$ Preparation. A large volume of 1-octadecene was degassed overnight at 110 °C under a vacuum of 2×10^{-2} Torr. The solution was then purged with N_2 and stored in a N_2-purged glovebox. A specific amount of $(TMS)_2S$ (listed below) was diluted with the dried ODE. This procedure was performed in a N_2-purged glovebox.

Chloride Additive Preparation. The required amount of $PbCl_2$ was added to a solution of oleylamine (98%, Sigma) in a 50 mL Schlenk flask and heated at 80 °C in a silicone oil bath until all of the $PbCl_2$ dissolves in the oleylamine. Should the $PbCl_2$ remain cloudy in solution, the Schlenk flask was purged with nitrogen, and oleic acid (90% tech. grade) was added dropwise to the solution to achieve a clear solution.

PbS CQD Synthesis. The $Pb(OA)_2$ stock solution was diluted in 15 mL of ODE at 100 °C to aid with dissolution, adding the required amount of OLA (0–1 mL), under a vacuum of 2×10^{-2} Torr, to achieve the required molar amount prior to $(TMS)_2S$ injection. For a given injection temperature, the $Pb(OA)_2$ solution was cooled to ~50 °C, followed by slow heating. As soon as the $Pb(OA)_2$ solution reaches the target injection temperature, the heating mantle was turned off and $(TMS)_2S$ solution was swiftly injected into the flask. One milliliter of the required molarity solution of $PbCl_2$ in oleylamine (typically, 0.3 mM) was injected into the flask when the whole mixture has cooled to 60 °C. The CQDs were allowed to cool to 30 °C during ~30 min before isolation. Accelerated cooling by removing the heating mantle, using a cold water bath or by injecting cold toluene, was found to be beneficial to monodispersity without affecting the target bandgap, if performed at a later stage of the cooling process. However, quenching is incompatible with chloride injection, since if it is done too late, it will have no visible effect on monodispersity, but if too early, it will negate the effect of chlorides. Acetone (distilled in glass, Caledon) was added to the Schlenk flask to precipitate the CQDs out of the solution. The cloudy solution was centrifuged and the supernatant was discarded. The wet CQDs were dissolved in toluene (distilled in glass, Caledon) and reprecipitated with acetone. The precipitation with acetone and dissolution in toluene was repeated at least twice.

Machine Learning Model. Machine learning is known to perform well in data interpolation, *i.e.*, in the regions where sufficient training data are available. Extrapolation to the regions far from the training data points is not reliable, yet the typical models provide predictions nevertheless. If these predictions are wrong, the model may choose not to explore the potentially promising regions of the input space, thus limiting the model's effectiveness in parameter optimization.

Bayesian models, in contrast, provide a confidence interval for their predictions, *i.e.*, a range of possible values instead of a single value. This broadens the search space, and if the predictions appear to be wrong upon experimental verification, the collected data are used to correct the model and improve its confidence near the added points.

Gaussian processes and Bayesian neural networks are the two major methods suitable for Bayesian optimization (*i.e.*, capable of estimating the confidence interval for their predictions). Gaussian processes are known to be too computationally intensive when the number of data points exceeds ~300. With this in mind, we chose to work with Bayesian neural networks.

A neural network with eight inputs (outdoor temperature, $PbOA_2$ volume, OLA volume, $(TMS)_2S$ injection temperature, $(TMS)_2S$ volume, ODE volume, molarity of $PbCl_2$ injected at max temperature, and $PbCl_2$ injected at 60 °C), two dense layers with 20 *elu* nodes each, and two outputs (bandgap and line width) was implemented using the Google TensorFlow library and Keras API.

To obtain the confidence interval, the dropout technique during both training and predictions was used.[21] This technique does not provide a rigorous way to assess the true value of uncertainty; thus, to encourage broader parameter space exploration, we opted for a larger value of dropout (10%), providing uncertainties comparable to the noise level in the experimental data (see SI Figure 1). The use of dropout requires a relatively wide network; we chose 20 nodes per layer, in which case a 10% dropout corresponds to two nodes missing. Levels of dropout below 5% would essentially result in a discrete switching between one and zero nodes.

Different model sizes (up to 100 nodes per layer and up to four layers) were tested in order to provide sufficient capacity in fitting the data. Dropout and L2 regularization parameters then were chosen based on conventional cross-validation to avoid overfitting and underfitting (SI Figure 8). Initially, a 50% validation split was used due to a high level of redundancy in experimental data. Further on, once the additional ML-driven experimental data points were collected, they were used for validation, while training on all pre-ML historic data. To make predictions regarding which additional experiments to perform, all data were combined into training and were trained using the model parameters chosen at the validation stage.

Additional visual analysis of the predictions was performed to analyze the onset of overfitting. In particular, plots of the predictions for a full swipe of input parameter values have a distinct shape in the 2D output space of wavelength and hwhm (see SI Figure 4). A too jagged shape of these prediction curves is a strong sign of overfitting. In addition, a robust model should result in the same predictions when retrained from scratch, especially in a region where enough training points are available; otherwise the model should be discarded and the degree of regularization increased. In this regard, we find that low regularization combined with early stopping, while providing seemingly good fits, is in fact too irreproducible.

The final model with two hidden layers, 20 nodes each, takes ~6 min on a modest desktop PC to be trained for one complete round of 15 000 epochs. Once the model is trained, data points (up to 1 000 000) in a subset of input space (either smaller or wider than that of initial experiments) are randomly generated using a uniform distribution, and the model predictions are recorded for each of them. The whole process takes less than a minute. This allows plotting a nearly continuous map, slicing through the dimensions of either input or output space.

Data Preparation. In the historical data, the line width was not explicitly measured, and only the peak/valley ratio was available. We

DOI: 10.1021/acsnano.9b03864
ACS Nano XXXX, XXX, XXX–XXX

thus devised a parameter proportional to exp(−peak/valley) that correlated well with the hwhm divided by the bandgap in samples where it was available, with the benefit that all CQD sizes would fall into one trend (see SI Figure 10). The only requirement on the polydispersity:hwhm relationship is that there should be a unique mapping between the two parameters used for model training, although they are not necessarily interchangeable. In our case, the chosen polydispersity metric is high, where hwhm is significantly worse than the expected homogeneous line width,[4,16] for any CQD size. On the other hand, it asymptotically reaches some value when peak/valley is high, reflecting the asymptotic behavior of the hwhm. In this manner, the prediction error penalizes inaccuracies in the low peak/valley range, while being less sensitive to the exact peak/valley value for nearly monodisperse samples. This polydispersity metric is used in all plots throughout.

Since polydispersity and wavelength are trained simultaneously, the total error comprises a sum of the two mean square errors. To increase the importance of one over the other, a scaling coefficient can be used. In the case of the Keras API, such coefficients cannot be explicitly defined, and we used scaling of one of the dimensions of the output data instead. In our case, the polydispersity metric is scaled down by a factor of 3, reducing its importance in fitting. The value of 3 is chosen based on plotting the noise distribution for multiple experiments using nominally the same parameter set (SI Figure 1)

The ambient lab temperature has a strong effect on the results of the synthesis, affecting the temperature of the injected precursors as well as the flask cooling rate. However, room temperature was not recorded in the historical data. We thus used seasonal outdoor temperature as a proxy to the temperature in the lab, although it does not necessarily predict the daily swings in temperature within each season.

We also noted that different supplier batches for PbO, OA, and $(TMS)_2S$ might produce consistently biased results, as well as $(TMS)_2S$ reactivity changes upon aging, even if stored in a glovebox. Unfortunately, there is no way to backtrack these parameters in the historical data, but monitoring them during future experiments might help with further optimization.

Several chloride additives were employed experimentally: $CdCl_2$, TBAC, $PbCl_2$, $AsCl_3$. They were included in ML training but not explicitly differentiated, as the number of such experiments was not sufficient to justify addition of an extra variable. $AsCl_3$ was found to be more active than other chlorides, requiring 5× lower molarity, which was compensated by appropriate rescaling in the ML training data. The use of $AsCl_3$ nevertheless did not provide any advantage in terms of improving hwhm, and the best results were observed with $PbCl_2$, which was used throughout all ML-driven experiments. The majority of the historic pre-ML data used $CdCl_2$.

No data oversampling/balancing was performed for freshly added experimental data points, thus representing lower certainty for experimental results that were performed fewer times, in the spirit of Bayesian learning.

Following the ML best practices, all input and output parameter values were normalized (rescaled) to the [0; 1] range prior to performing the model training, in order to ensure more robust convergence. As a result, all output values require scaling back for converting them to true values.

Raw data as well as the fully trained model are available in the Supporting Information.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acsnano.9b03864.

Model learning with different amounts of training data, overfitting and underfitting prevention in the presence of experimental noise, effect of Cl and OLA on polydispersity, XRD and TEM of representative CQD batches, comparison to published absorption spectra, relation between the polydispersity metric and hwhm (PDF)

Raw data and fully trained model (ZIP)

AUTHOR INFORMATION

Corresponding Author

*E-mail: ted.sargent@utoronto.ca.

ORCID

Oleksandr Voznyy: 0000-0002-8656-5074
James Z. Fan: 0000-0002-1594-865X
Olivier Ouellette: 0000-0001-5708-5058
Petar Todorović: 0000-0002-2838-876X
Edward H. Sargent: 0000-0003-0396-6495

Present Address

#Department of Mechanical Engineering, IIT Bombay, Powai, Mumbai 400 076, Maharashtra, India.

Author Contributions

†O. Voznyy, L. Levina, and J. Z. Fan contributed equally.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the Ontario Research Fund Research Excellence Program and by the Natural Sciences and Engineering Research Council (NSERC) of Canada. The authors thank E. Palmiano, R. Wolowiec, D. Kopilovic, and C. S. Tan for their help during the course of this study.

REFERENCES

(1) Hines, M. A.; Scholes, G. D. Colloidal PbS Nanocrystals with Size-Tunable Near-Infrared Emission: Observation of Post-Synthesis Self-Narrowing of the Particle Size Distribution. *Adv. Mater.* **2003**, *15*, 1844−1849.

(2) Cademartiri, L.; Bertolotti, J.; Sapienza, R.; Wiersma, D. S.; von Freymann, G.; Ozin, G. A. Multigram Scale, Solventless, and Diffusion-Controlled Route to Highly Monodisperse PbS Nanocrystals. *J. Phys. Chem. B* **2006**, *110*, 671−673.

(3) Moreels, I.; Justo, Y.; De Geyter, B.; Haustraete, K.; Martins, J. C. J. C.; Hens, Z. Size-Tunable, Bright, and Stable PbS Quantum Dots: A Surface Chemistry Study. *ACS Nano* **2011**, *5*, 2004−2012.

(4) Weidman, M. C.; Beck, M. E.; Hoffman, R. S.; Prins, F.; Tisdale, W. A. Monodisperse, Air-Stable PbS Nanocrystals *via* Precursor Stoichiometry Control. *ACS Nano* **2014**, *8*, 6363−6371.

(5) Hendricks, M. P.; Campos, M. P.; Cleveland, G. T.; Jen-La Plante, I.; Owen, J. S. A Tunable Library of Substituted Thiourea Precursors to Metal Sulfide Nanocrystals. *Science* **2015**, *348*, 1226−1230.

(6) LaMer, V. K.; Dinegar, R. H. Theory, Production and Mechanism of Formation of Monodispersed Hydrosols. *J. Am. Chem. Soc.* **1950**, *72*, 4847−4854.

(7) Zhang, J.; Crisp, R. W.; Gao, J.; Kroupa, D. M.; Beard, M. C.; Luther, J. M. Synthetic Conditions for High-Accuracy Size Control of PbS Quantum Dots. *J. Phys. Chem. Lett.* **2015**, *6*, 1830−1833.

(8) Fu, H.; Tsang, S.-W.; Zhang, Y.; Ouyang, J.; Lu, J.; Yu, K.; Tao, Y. Impact of the Growth Conditions of Colloidal PbS Nanocrystals on Photovoltaic Device Performance. *Chem. Mater.* **2011**, *23*, 1805−1810.

(9) Ip, A. H.; Thon, S. M.; Hoogland, S.; Voznyy, O.; Zhitomirsky, D.; Debnath, R.; Levina, L.; Rollny, L. R.; Carey, G. H.; Fischer, A.; Kemp, K. W.; Kramer, I. J.; Ning, X.; Labelle, A. J.; Chou, K. W.; Amassian, A.; Sargent, E. H. Hybrid Passivated Colloidal Quantum Dot Solids. *Nat. Nanotechnol.* **2012**, *7*, 577−582.

DOI: 10.1021/acsnano.9b03864
ACS Nano XXXX, XXX, XXX−XXX

(10) Rempel, J. Y.; Bawendi, M. G.; Jensen, K. F. Insights into the Kinetics of Semiconductor Nanocrystal Nucleation and Growth. *J. Am. Chem. Soc.* **2009**, *131*, 4479−4489.
(11) Cao, B.; Adutwum, L. A.; Oliynyk, A. O.; Luber, E. J.; Olsen, B. C.; Mar, A.; Buriak, J. M. How To Optimize Materials and Devices *via* Design of Experiments and Machine Learning: Demonstration Using Organic Photovoltaics. *ACS Nano* **2018**, *12*, 7434−7444.
(12) Dewancker, I.; McCourt, M.; Clark, S.; Hayes, P.; Johnson, A.; Ke, G. A Stratified Analysis of Bayesian Optimization Methods. 1603.09441.arXiv, **2016**, https://arxiv.org/abs/1603.09441 (accessed September 18, 2019).
(13) Bezinge, L.; Maceiczyk, R. M.; Lignos, I.; Kovalenko, M. V.; DeMello, A. J. Pick a Color MARIA: Adaptive Sampling Enables the Rapid Identification of Complex Perovskite Nanocrystal Compositions with Defined Emission Characteristics. *ACS Appl. Mater. Interfaces* **2018**, *10*, 18869−18878.
(14) Voznyy, O.; Levina, L.; Fan, F.; Walters, G.; Fan, J. Z.; Kiani, A.; Ip, A. H.; Thon, S. M.; Proppe, A. H.; Liu, M.; Sargent, E. H. Origins of Stokes Shift in PbS Nanocrystals. *Nano Lett.* **2017**, *17*, 7191−7195.
(15) Thon, S. M.; Ip, A. H.; Voznyy, O.; Levina, L.; Kemp, K. W.; Carey, G. H.; Masala, S.; Sargent, E. H. Role of Bond Adaptability in the Passivation of Colloidal Quantum Dot Solids. *ACS Nano* **2013**, *7*, 7680−7688.
(16) Peterson, J. J.; Krauss, T. D. Fluorescence Spectroscopy of Single Lead Sulfide Quantum Dots. *Nano Lett.* **2006**, *6*, 510−514.
(17) Whitham, K.; Yang, J.; Savitzky, B. H.; Kourkoutis, L. F.; Wise, F.; Hanrath, T. Charge Transport and Localization in Atomically Coherent Quantum Dot Solids. *Nat. Mater.* **2016**, *15*, 557−563.
(18) Cui, J.; Beyler, A. P.; Coropceanu, I.; Cleary, L.; Avila, T. R.; Chen, Y.; Cordero, J. M.; Heathcote, S. L.; Harris, D. K.; Chen, O.; Cao, J.; Bawendi, M. G. Evolution of the Single-Nanocrystal Photoluminescence Linewidth with Size and Shell: Implications for Exciton-Phonon Coupling and the Optimization of Spectral Linewidths. *Nano Lett.* **2016**, *16*, 289−296.
(19) An, J. M.; Franceschetti, A.; Dudiy, S. V.; Zunger, A. The Peculiar Electronic Structure of PbSe Quantum Dots. *Nano Lett.* **2006**, *6*, 2728−2735.
(20) Razgoniaeva, N.; Yang, M.; Garrett, P.; Kholmicheva, N.; Moroz, P.; Eckard, H.; Royo Romero, L.; Porotnikov, D.; Khon, D.; Zamkov, M. Just Add Ligands: Self-Sustained Size Focusing of Colloidal Semiconductor Nanocrystals. *Chem. Mater.* **2018**, *30*, 1391−1398.
(21) Gal, Y.; Ghahramani, Z. Bayesian Convolutional Neural Networks with Bernoulli Approximate Variational Inference. **2015**, 1506.02158.arXiv. https://arxiv.org/abs/1506.02158 (accessed September 18, 2019).

DOI: 10.1021/acsnano.9b03864
ACS Nano XXXX, XXX, XXX−XXX