Memory in quantum dot blinking

Roberto N. Muñoz [1,*] Laszlo Frazer [2] Gangcheng Yuan,[2] Paul Mulvaney [3] Felix A. Pollock,[4] and Kavan Modi [1]

[1]*ARC Centre of Excellence in Exciton Science and School of Physics & Astronomy, Monash University, Clayton, Victoria 3800, Australia*
[2]*ARC Centre of Excellence in Exciton Science and School of Chemistry, Monash University, Clayton, VIC 3800, Australia*
[3]*ARC Centre of Excellence in Exciton Science, School of Chemistry, University of Melbourne, Parkville, VIC 3010, Australia*
[4]*School of Physics & Astronomy, Monash University, Clayton, Victoria 3800, Australia*

The photoluminescence intermittency (blinking) of quantum dots is interesting because it is an easily measured quantum process whose transition statistics cannot be explained by Fermi's golden rule. Commonly, the transition statistics are power-law distributed, implying that quantum dots possess at least trivial memories. By investigating the temporal correlations in the blinking data, we demonstrate with high statistical confidence that there is nontrivial memory between the on and off brightness duration data of blinking quantum dots. We define nontrivial memory to be statistical complexity greater than one. We show that this memory cannot be discovered using the transition distribution. We show by simulation that this memory does not arise from standard data manipulations. Finally, we conclude that at least three physical mechanisms can explain the measured nontrivial memory: (1) storage of state information in the chemical structure of a quantum dot; (2) the existence of more than two intensity levels in a quantum dot; and (3) the overlap in the intensity distributions of the quantum dot states, which arises from fundamental photon statistics.

I. INTRODUCTION

Quantum dots (QDs) are nanoscale semiconducting crystals [1–3] with a major role in developing energy-efficient technology [4,5]. When used as light-emitting diodes, QDs can be synthesized to emit light at precise colors which has resulted in the production of high-quality displays [6]. An attractive advantage in using QDs for everyday and commercial use is that they outperform conventional LEDs in terms of output and cost efficiency [7], allowing spaces to be lit for less.

Despite these promises, the integration of QDs into scalable lighting applications is limited by intermittent photoluminescence known as blinking [8–11]. The blinking phenomenon is observed over a variety of QDs with different compositions [12,13], and results in periods where an excited QD enters an *off* state and no longer emits photons [14]. The intermittent lack of photons makes lighting less efficient [12,15], and hence mitigating blinking is widely studied.

To mitigate blinking, it is necessary to understand the physics and chemistry behind the phenomena. Blinking is generally characterized from data by employing statistical methods like duration distributions [9,16], power spectra [17], or autocorrelations [18–21]. A number of chemical and physical pathways have been proposed to contribute to blinking using these methods [8,13,14,22–24]. A possible way to better understand the blinking process is to look for correlations in the blinking patterns, i.e., characterize the blinking using information linking the past and future blinking behavior in a time series. Such correlations are often referred to as *memory*

[25]. Information theory reveals that the memory in a data sequence requires special tools to infer it, and is inaccessible via the aforementioned statistical methods [26].

In this paper, we show that the on and off duration sequences in a blinking quantum dot indeed have memory. We do this by using methods from computational mechanics which infer memory from data [27–29]. We use hypothesis testing to show a model with trivial memory is a significantly worse explanation of the data than a model with nontrivial memory. Importantly, the model with trivial memory is rejected, even though it is an exact fit to the distribution of blinks.

We then relate the observation of blinking memory to a number of processes which may be acting within the QDs in our analysis. Our methods are applicable to a wide range of chemical systems, and function as a flexible and scalable stepping point for future memory-based analyses. We aim that the ability for memory to be associated with specific chemical processes can ultimately serve as a diagnostic tool for the improvement of QD synthesis techniques, and overall output quality.

The manuscript is structured as follows: Sec. II A is a brief description of the QD blinking data used in this study, in addition to the elementary processing needed to separate the brightness states. Section II B covers the on and off state durations and the statistics which they exhibit. It focuses on the power-law behavior observed in some blinking QDs, and how it relates to searches of memory. We then look at how memory is described in Sec. III A and how it can be inferred from on and off duration data in Sec. III B. We also cover how hypothesis testing can compare different models of memory in Sec. III C, before applying this methodology to the on and off duration blinking data. The main results are shown in

─────────
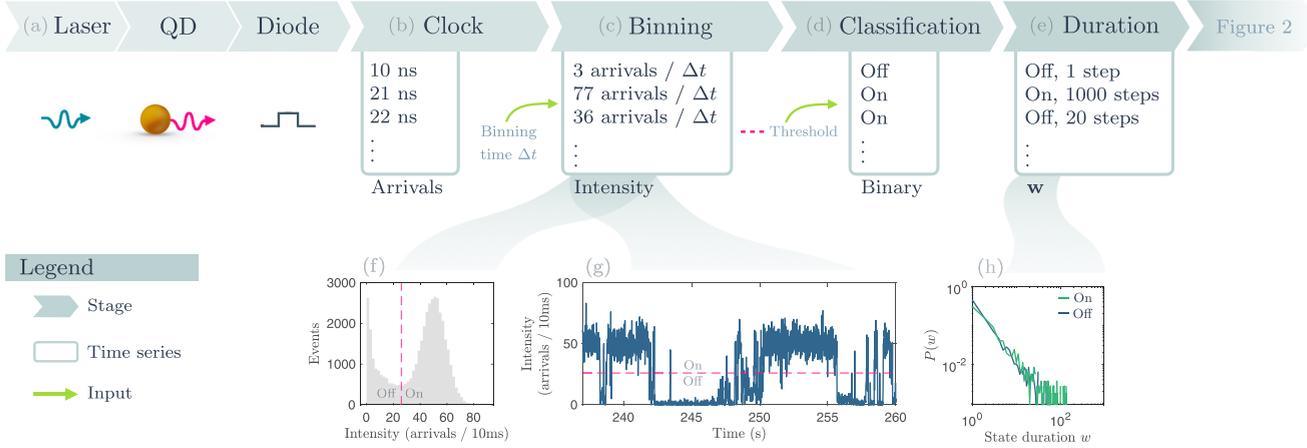*roberto.munoz@monash.edu

014127-1

FIG. 1. A flow chart representing data collection and processing. (a) A laser provides a beam of light which excites quantum dots that emit luminescence. The luminescence photons are converted to electrical pulses using an avalanche photodiode. (b) A nanosecond resolution clock records the time when electrical pulses are received. (c) The data is divided into time bins with width Δt, chosen to maximize contrast between the peaks in the histogram of intensities (f). The number of photon arrival times in each bin is computed. (d) Each time bin is classified by using an intensity threshold I_{thresh} to determine if the quantum dot is on or off during the time bin. (e) The duration of on and off states is computed. (f)–(h) Summary data taken from QD4@200 nW. (f) Intensity histogram obtained by binning arrival times over 10 ms windows. Dashed line marks the intensity threshold, taken to be at the minima between the on and off peaks. (g) Portion of the intensity trace of the blinking QD. Solid line indicates the recorded intensity over time with clear indications of the blinking intermittency. (h) Distribution of state durations graphed on a log-log scale. Data exhibits linear behavior characteristic of power-law-like statistics

Sec. III D, followed by a discussion on potential sources for the observed memory in Sec. IV, with simulated examples.

II. QUANTUM DOT DATA AND TREATMENTS

A. Data collection and processing

We start by covering only the essential details of QD synthesis and measurement necessary for this study. Further elaboration may be consulted from Ref. [8], with Supplemental Material showing summary plots of all quantum dot data recordings used in this study found in Ref. [30].

Our study used core-shell CdSe/Cd$_x$Zn$_{1-x}$S quantum dots, previously synthesised [31,32] and measured in Ref. [8]. A total of 7 identically prepared quantum dots were each excited with a pulsed laser diode, and measured over excitation powers ranging from 100 nW to 600 nW. The photoluminescence of single QDs was collected using a microscope onto an avalanche photodiode detector. A nanosecond-resolution clock was then used to record the times of photon arrivals coming from an excited QD. We binned the photon arrival time measurements into 10ms intervals to convert them into a sequence of intensities

$$\mathbf{I} = (I_1, I_2, \ldots, I_k, \ldots), \tag{1}$$

as shown in Fig. 1(c). An example of the intensity histogram for a blinking quantum dot is shown in Fig. 1(f).

Most blinking dots generally feature two prominent peaks in intensity distributions: one corresponding to a *bright/on* state and the other corresponding to a *dark/off* state. In the bright state, photons arrive in the detector with a rate λ_{On}, while the dark state has an arrival rate λ_{Off} that is governed by background. Quantum dots with high emitting rates that spend the majority of the measured time in the on state, typically exhibit a large amount of bright counts relative to

dark. This is observed by the relative heights of the two peaks in the distribution [Fig. 1(f)], in addition to high contrast between the peaks. We found binning photon arrival times over 10ms windows gave the best contrast between intensity peaks. When investigating the nature of quantum dot blinking, being able to classify a series of intensity data into bright and dark states is a crucial step for understanding the underlying chemical processes. For each quantum dot, this was performed by selecting an intensity threshold I_{thresh} located at the minima between the bright and dark intensity peaks. For each intensity outcome I_k in the measured sequence $\mathbf{I}$ [Eq. (1)], we classified them into on and off states using the threshold, $I_{\text{Off}} \leqslant I_{\text{thresh}} < I_{\text{On}}$, to form a sequence of binary on and off classifications [Fig. 1(d)].

B. State durations and power-laws

The separation of intensity data into binary on and off states is a common way to explore blinking quantum dots. For example, by scanning through the classification sequence and recording the length of consecutive on and off outcomes, one obtains an alternating sequence of on and off state durations

$$\mathbf{w} = \left(w_1^{\text{On}}, w_2^{\text{Off}}, \ldots, w_k^b, \ldots \right) \; : \; b \in \{\text{On}, \text{Off}\}, \tag{2}$$

shown in Fig. 1(e). The measured sequences of on and off state durations described by Eq. (2) are the primary objects of this study. State durations, where the amount of time w^b a quantum dot can spends in brightness state b before blinking can be modeled as a pair of probability distributions $P_b(w^b)$. These distributions were regarded to be power-law distributed [9,12,15,16]

$$P_b(w^b) \propto (w^b)^{-m_b}, \tag{3}$$

where m_b is the power-law constant, identified by linear behavior when plotted on a log-log scale [Fig. 1(h)]. The most remarkable feature regarding power-law state durations and blinking behavior, is the range in state durations over orders of magnitude [33]. This means that quantum dots have been observed to blink at timescales orders of magnitudes in difference, from milliseconds to hours [16,34]. Recent studies over a wider range of QD compositions and synthesis techniques have demonstrated the existence of exponential, multiexponential, and quasiexponential distributions for duration statistics [8,35]. Identifying the type of distribution present in blinking data has allowed experimenters to relate them to underlying chemical processes, e.g., exponential $P_b(w^b) \propto \exp(-Cw)$ and power-law durations related to two different blinking mechanisms in Ref. [8].

Fermi's golden rule says that transitions between eigenstates of the Hamiltonian are exponentially distributed. This suggests that, for blinks which are not exponentially distributed, the on and off states are actually collections of multiple eigenstates. If those eigenstates correspond to different chemical structures, then they may tend to occur in certain sequences. For example, changes to surface bonding can enable changes one layer beneath the surface [36–38]. Therefore, there is reason to suspect that additional information can be inferred from the blinking sequence **w**, than just the distributions from Eq. (3) alone.

Notably, the fact that state duration distributions can often be described by power-law statistics is intriguing in its own right. For example, in statistical mechanics, power-law distributions arise for systems at critical points such as during phase transitions [39]. The appearance of power-law statistics in other fields is similarly attributed to systems of interest whose dynamics are suspected to reside at criticality [40,41], or that are scale invariant [42]. For quantum dots, power-law distributed on and off duration times are related to phenomena such as ageing and nonergodicity [43–45] in the sequences of intensity recordings, which may stem from a variety of physical reasons. As such, some time-averaged quantities derived from the intensity process are prone to being dependent upon the total observation time when the durations are power-law distributed [43–46]. As a result, considering the on and off durations as being drawn from a *stationary* power-law distribution [Eq. (3)] is a simplifying assumption made in this manuscript. Another central idea surrounding systems with power-law statistics is that they possess high degrees of *memory* [47]. That is, systems which depend on correlations over multiple points of time in the past [25]. While we unpack the notion of memory in the next section, the main takeaway here is that memory arises in a multitude of complex systems over a breadth of research fields. Some examples range from open quantum systems [48,49] to the firing patterns of neurons in the brain [50,51]. Finding equations which predict the future evolution of systems with memory is difficult, since the past behavior of the system must be taken into account [52,53].

If observing power-law distributions alone truly is the "litmus test" for memory, then one would expect there to be multitime correlations in the power-law distributed on and off state durations of Eq. (2). Ultimately, histogramming any temporal data on a log-log scale and observing a linear trend would be sufficient to mark the presence of memory, and no

other additional tools would be needed. However, this is not the case. There is a body of research which counters the idea that power-law distributions imply a deep structure and high memory. Power-law distributions can arise from an ensemble of independent, random, and unstructured processes, none of which at the individual level contain multitime correlations [54]. Notable examples include, combining multiple exponential distributions together [55,56], and even an ensemble of monkeys on typewriters [57]. Outside of memory, the observation of power-law distributions is also not always sufficient to determine the underlying mechanics that cause it [58]. Even using other statistical treatments such as autocorrelation [18–21] or power spectra [17] are shown to be insufficient to reveal memory in some cases [26]. Ultimately, the ability to infer the presence of memory in the state durations of blinking quantum dots depends on the use of specialised tools.

III. NON-MARKOVIANITY, COMPUTATIONAL MECHANICS

We aim to infer the memory present in the on and off state duration data that is inconclusive via the observation of power-law distributed on and off durations alone. We describe the specialised tools necessary for this task, in addition to the mathematical framework they entail. We start by defining memory as the amount of past blinking information required to predict the future behavior of the data (*non-Markovianity*) [27], and how it relates to the length of past observation sequences (*Markov order*) [53]. We then describe how models can efficiently characterize memory. Next, we explain how to construct these models from data, and which model corresponds to the hypothesis with trivial memory. We then go on to outline the hypothesis testing methods used to reject the null model. Finally, we find that given the data, we can reject models with trivial memory compared to ones which include memory in the data.

A. Time series and prediction

Regardless of the physical system, inferring the memory of a process requires measurement data to be described as a *time series* [59]. This means that at each time step k the system is measured and the measurement outcomes r_k at each step form a sequence $\mathbf{r} = (r_1, \ldots, r_k, \ldots)$. For the case of a blinking quantum dot, general examples of time series could include the sequence of intensity measurements [Fig. 1(c)], but in our work we consider only the sequence of brightness state durations [Fig. 1(e) and Eq. (2)]. Each element of $\mathbf{r}$ is drawn from the set of measurement outcomes, or *alphabet* $\mathcal{A}$, such that $r_k \in \mathcal{A}$. For most applications, the time series $\mathbf{r}$ is considered over discrete time steps rather than continuous, in addition to considering the set of measurement outcomes $\mathcal{A}$ to be finite. When the occurrence of a measurement outcome at a given time step is random, the system generating the time series is a stochastic process [60,61]. The sequential measurements of state durations **w** mentioned in Eq. (2) are an example of a stochastic process since they can in general be modeled as random outcomes of a probability distribution. While the exact distribution can vary depending on the physical system being considered, we will assume power-law

distributions following Eq. (3) in this manuscript given the typical observation of multiscale durations. It is important to note however, that although stochastic, measurement outcomes may not appear in a completely independent manner. The probability of seeing a particular outcome in the future may strongly depend on the sequence of outcomes observed in the past. This concept defines the notion of *memory* when referring to stochastic time series. Being explicitly linked to past/future correlations, memory plays a central role in the prediction of the behavior of stochastic processes [25,53,61].

Relative to an arbitrary time step k, we denote the future and the past of a time series as $\mathbf{r} = (\bar{\mathrm{r}}, \vec{r})$. Explicitly the past and the future are $\bar{\mathrm{r}} = (\ldots, r_{k-2}, r_{k-1})$ and $\vec{r} = (r_k, r_{k+1}, \ldots)$, respectively. Processes which are memoryless follow the *Markov property*

$$P(r_k \mid \bar{\mathrm{r}}) = P(r_k \mid r_{k-1}), \tag{4}$$

which states that future evolution depends solely on the current outcome and not on any of the past observations [25,61]. In general, however, prediction of the immediate future outcome r_k, requires knowledge of the past L outcomes $\bar{\mathrm{r}}_L := (r_{k-L}, \ldots, r_{k-2}, r_{k-1})$. Systems where this is true have memory, and are known as *non-Markovian*. Formally stated, non-Markovian systems satisfy

$$P(r_k \mid \bar{\mathrm{r}}_L) \neq P(r_k \mid r_{k-1}). \tag{5}$$

The length L for when outcome probabilities become independent of past observations of length $> L$ is known as the Markov order [61,62]. For the sequence of state duration data $\mathbf{w}$, this means that predicting the amount of time a QD remains in either the on and off state before blinking is dependent on how long it spent in the on and off state at multiple instances in the past. Processes which follow Eq. (4) are Markov processes and have $L = 1$, which we will show is related to a memoryless hypothesis for the blinking.

B. Memory modeled by HMMs

We aim to develop and perform a hypothesis comparison framework to compare between memoryless and non-Markovian hypotheses. To do this, we need two models which represent the Markovian and a non-Markovian hypotheses. In this section we outline how to construct these models which represent memory in the data.

For non-Markovian systems, Eq. (5) suggests that memory is related to the prediction of a process' future behavior based off past observations, and is proportional to Markov order. Although the difficulty in predicting the future behavior of non-Markovian processes rises exponentially with Markov order, there are efficient ways construct models that represent the underlying memory. For example, not all past observation sequences may result in distinct future behaviors, and hence not every possible past sequence must be remembered by a non-Markovian process.

This was the central idea in work by Crutchfield and Young [27,63], whereby distinct past sequences that caused statistically equivalent future behavior were grouped together. Formally, two distinct sequences of past observations $\bar{\mathrm{r}}_L$ and $\bar{\mathrm{r}}'_L$ of length L belong to the same grouping $S_i \in \mathcal{S}$, if the probability of observing a specific future $\vec{r}$ given $\bar{\mathrm{r}}_L$ or $\bar{\mathrm{r}}'_L$ is

the *same*. That is,

$$\bar{\mathrm{r}}_L \sim_\epsilon \bar{\mathrm{r}}'_L \quad \text{if} \quad P(\vec{r} \mid \bar{\mathrm{r}}_L) = P(\vec{r} \mid \bar{\mathrm{r}}'_L), \tag{6}$$

where $\sim_\epsilon$ indicates that two past sequences correspond to the same partition. The entire collection $\mathcal{S}$ of all the partitions S_i are known as the *causal states* of the process.

Due to the nature of their construction, new causal states are only created when necessary and hence relate to the *minimal* amount of memory a non-Markovian process needs to store to determine future behavior. In the context of memory, the causal states are the hidden states that a non-Markovian system can be in.

The Shannon entropy [64,65] over the causal states quantifies the minimal number of bits of information required to optimally predict the future behavior of the process. The measure introduced in Ref. [27] is called the *statistical complexity*,

$$C_\mu := H[\mathcal{S}] = -\sum_i P(S_i) \log P(S_i), \tag{7}$$

and it is the memory of a non-Markovian process.

When the causal states $\mathcal{S}$ are known, the probabilities to transition from one causal state to another $\mathcal{T}$ while observing an outcome r_k can be estimated from the data. In general, a collection of states, transitions, and observations is a *hidden Markov model* (HMM) [60,61,65] and allows the causal states of a non-Markovian process in Eq. (6) to be represented by a directed graph [27,62]. When the vertices are the causal states, the edges are the transition probabilities between causal states, and labels on the edges are the observed outcomes, the set $\{\mathcal{S}, \mathcal{T}, \mathcal{A}\}$ is a special class of HMM called an ϵ-*machine* [27,63]. The ϵ-machine both represents the memory and models future behavior of a non-Markovian time series [Fig. 2(g)]. For the on and off duration sequences of quantum dots, the causal states have chemical and physical explanations such as the correlations obtained by having multiple brightness states, or varying emitting rates. We examine these examples and more in Sec. IV.

We infer the causal states, ϵ-machines, and complexities C_μ of the state duration sequence data $\mathbf{w}$ by using the *causal state splitting reconstruction* (CSSR) algorithm [28,66,67]. The ϵ-machine output by CSSR is the model will represent non-Markovian component for the model comparison testing in our study. We will refer to this alternative model as the ϵ-*machine model* (ϵMM) hypothesis.

While we refer details of the algorithm to Ref. [28] and Appendix A, the successful operation of CSSR relies on some key attributes related to the input data. The first, is the algorithm requires as input the maximum length $L_{\max}$ of past sequences to consider for the grouping of past observations in Eq. (6). If the non-Markovian memory contained within the time series has an intrinsic Markov order L, then choosing $L_{\max} < L$ results in poor prediction due to the inferred ϵMM not capturing the long-memory correlations in the data. However, CSSR will still produce a model that is consistent with the non-Markovian correlations up to order L. Given sufficient data, choosing $L_{\max} \geqslant L$ guarantees convergence on the true extent of non-Markovian correlations.

The second is that given the finite length of data available, there is an upper limit on values of $L_{\max}$ above which the resulting conditional probabilities $P(\vec{r} \mid \bar{\mathrm{r}}_L)$ estimated in
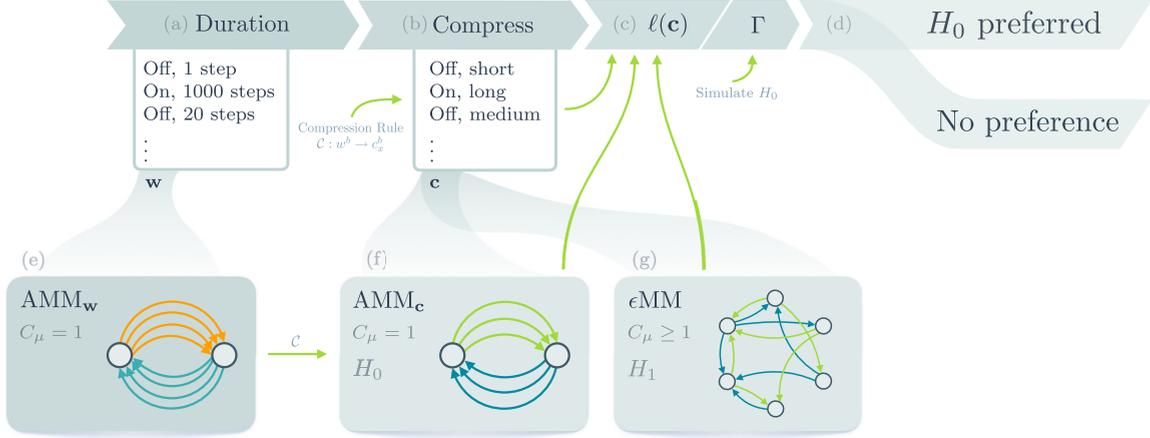
FIG. 2. Flow chart representing the data compression, construction of the AMM and ϵMM models, and model comparison. (a) The duration of on and off states is converted to a sequence of reduced measurement outcomes via the compression rule $\mathcal{C}$ in Eq. (12) to meet the constraints described in Sec. III B. (e) The alternating Markov model for a time series of on and off state durations **w**. Each gold (right-pointing) arrow is associated with the probability for a specific on duration to be observed at the immediate future time step in the sequence. Blue (left pointing) arrows represent the same for an off duration. The nodes are the causal states. (b) Time series of compressed state durations, ready for the hypothesis testing stage. (f) The AMM fit to compressed data **c**. Each green (right pointing) arrow represents the probability for the coarse on duration to be observed at a future time step in the sequence. Blue (left pointing) arrows are the same for a coarse off duration. This model is the result of performing the compression $\mathcal{C}$, and results in the same alternating structure, but with fewer measurement outcomes (arrows) as desired. (g) An example of an ϵMM inferred from the compressed data by the CSSR algorithm up to a Markov order $L_{\max}$. The graph structure shown here is an example of what is representative in QDs with memory. In general, ϵMMs inferred from non-Markovian data have more nodes and are hence more complex than the AMM. (c) Hypothesis comparison stage consisting of the negative log-likelihood $\ell(\mathbf{c})$ and model preference threshold Γ. The $\ell(\mathbf{c})$ takes the compressed data **c**, AMM, and ϵMM models as inputs. The threshold is computed by simulating the AMM. (d) Hypothesis comparison decision either shows preference for the AMM, or fails to draw preference between the AMM and ϵMM. For each time-series data, $\ell(\mathbf{c})$ is compared to the rejection threshold Γ as shown in Eq. (9). Preference over the ϵMM hypothesis over the AMM indicates significant non-Markovian correlations in the blinking data.

Eq. (6) will be prone to severe under-sampling. This results in the algorithm creating new causal states for every string of length $L_{\max}$ it encounters. To avoid this issue, the range of permissible past sequence lengths $L_{\max}$ given data of length N is bounded by

$$L_{\max} \leqslant \frac{\log N}{\log |\mathcal{A}|}, \qquad (8)$$

where $|\mathcal{A}|$ are the number of unique measurement outcomes in the time-series data [28,68]. Generally speaking, the length of data N suitable for CSSR is entirely dependent on the alphabet size and maximum desired Markov order.

For the memoryless or Markovian hypothesis, the corresponding model can be deduced in simple cases. This can be done to anticipate the memory structure of the time series of alternating QD state durations in Eq. (2), should a blinking quantum dot store no information about its past behavior. This is true if on and off state durations w are drawn independently from their respective distributions $P_b(w^b)$. The probability of observing the immediate future state duration w_k^b, would only need to know which brightness state b the state duration at the previous time step w_{k-1}^b belonged to. With longer past sequences having no bearing on future behavior, a blinking quantum dot with no correlations has an intrinsic Markov order 1.

The minimal memory model therefore consists of two causal states $|\mathcal{S}| = 2$, one storing all past sequences with the most recent state durations coming from the on state $\bar{\mathbf{w}} =$

$(\ldots, w_{k-2}^{\mathrm{Off}}, w_{k-1}^{\mathrm{On}})$, and the other for past sequences ending in a state duration from the off state $\bar{\mathbf{w}}' = (\ldots, w_{k-2}^{\mathrm{On}}, w_{k-1}^{\mathrm{Off}})$. Any ϵ-machine that alternates between two causal states always has a memory of $C_\mu = 1$, and is shown in Fig. 2(e). We refer to the alternating model with $C_\mu = 1$ as the *alternating Markov model* (AMM), which can be constructed from alternating data **w** by estimating the empirical distributions $\hat{P}_{\mathrm{On}}(w^{\mathrm{On}}) = \nu(w^{\mathrm{On}})/N_{\mathrm{On}}$ and $\hat{P}_{\mathrm{Off}}(w^{\mathrm{Off}}) = \nu(w^{\mathrm{Off}})/N_{\mathrm{Off}}$ from time series. The $\nu(w^b)$ term refers to the frequency counts for the outcome w^b in the data, and $N_b : b \in \{\mathrm{On}, \mathrm{Off}\}$ are the total number of on and off symbols in the data respectively. The probabilities from $\hat{P}_b(w^b)$ were used as the transition probabilities for the AMM, arranged following the structure in Fig. 2(e). We note that although an AMM and ϵMM when inferred at $L_{\max} = 1$ may be the same, this is not always true in general [28].

C. Hypothesis testing

Complexity, C_μ is the memory of an underlying process, with large values associated with a highly non-Markovian process [Fig. 2(g)]. It is worth noting that while two HMMs can be statistically distinct in the future behaviors they produce, they may share the same C_μ and/or produce the same distribution over outcomes [Eq. (3)]. Figure 3 demonstrates an example of the latter. This highlights the need for being able to statistically compare models when given data. Conducting statistical comparisons that use data to favor models requires
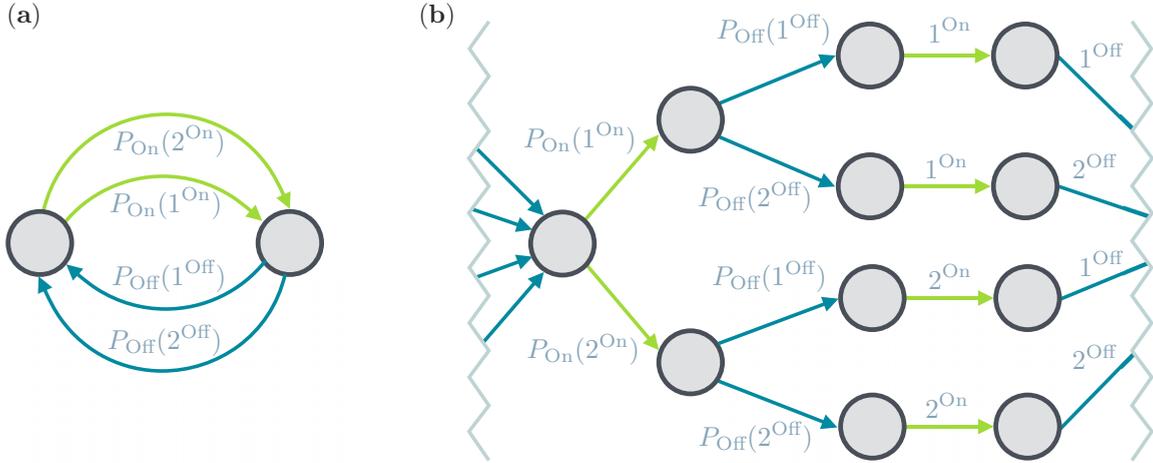
FIG. 3. Examples of two HMMs that have different memory and which result in the same overall state duration distribution $P_b(w^b)$. The HMMs can be "read" as follows: Pick a starting causal state (gray node), and follow any outgoing arrow to another causal state. This represents a measurement outcome which occurs with the labeled probability. Green (light) arrows represent on measurement outcomes, blue (dark) arrows represent off outcomes. If a causal state has only one outgoing arrow, then the labeled measurement outcome occurs with probability 1. Repeat these steps from the next causal state and so on, recording outcomes at each transition. (a) Example of an alternating Markov model (AMM) with two possible on and two possible off durations, 1 and 2. (b) Example of a HMM with the same possible on and off outcomes and probabilities as the AMM, but past sequences of length 2 are cloned. Starting from the leftmost state, an on duration followed by an off duration are sampled from the distributions $P_{\mathrm{On}}(w^{\mathrm{On}})$ and $P_{\mathrm{Off}}(w^{\mathrm{Off}})$. The following two transitions clone the outcomes and then return back to the leftmost state to begin the process again. The on and off duration distributions result the same graph for both models despite the HMM on the right having higher memory. The discrete probability distributions $P_b(w^b)$ are left unspecified for a purpose. They could be power-law, exponential, or something else entirely, and the overall structure of the two HMMs would not change. This also highlights that the specific distribution of the state durations does not indicate the presence or lack of memory.

data $\mathbf{r}$, a null hypothesis H_0, and an alternate hypothesis H_1. Unlike some other statistical scenarios, examining memory models in alternating sequences of waiting times does not have a built-in or obvious H_0, nor does it have an obvious complementary hypothesis to the null; hence we used a pair of behaviorally motivated models to conduct model comparison rather than rejection.

To show that quantum dots have nontrivial memory in the blinking durations, our null hypothesis H_0 is the model with trivial memory, which is the AMM from Sec. III B. This was motivated by the fact that there can be no other ϵ-machine with $C_\mu < 1$ for a memoryless, alternating time series. However, there are an infinite number of possible HMMs to use as the alternate model for comparison. Hence, selecting the ϵ-machine– the model generated by CSSR, has the advantage that it represents the HMM with the *minimal* degrees of freedom needed to predict a process with memory [28]. That is, there is no unifilar HMM that is more complex than the AMM but less complex than the ϵMM which to serve as an alternative. Hence choosing the ϵMM forms a suitably motivated alternative model, as the potentially infinite space of HMM alternatives is narrowed down. As such, by using minimal alternatives, we do not consider model comparison tests which penalise for overfitted models since there is no general way to reduce an ϵ-machine to a simpler model without loss of information. Given the AMM as the null model H_{AMM} [Fig. 2(f)], the ϵMM with $C_\mu \geqslant 1$ as the alternative model $H_{\epsilon\mathrm{MM}}$ [Fig. 2(g)], and a general time series $\mathbf{r}$, the *negative log-likelihood ratio* (nLLR) compares the two non-complementary models and determines which of the two is

preferred by the data via the rule

$$\ell(\mathbf{r}) < \Gamma \;\rightarrow\; H_{\epsilon\mathrm{MM}} \text{ preferred over } H_{\mathrm{AMM}},$$
$$\ell(\mathbf{r}) \geqslant \Gamma \;\rightarrow\; \text{fail to prefer } H_{\mathrm{AMM}}, \tag{9}$$

where

$$\ell(\mathbf{r}) = -2\log\frac{P(\mathbf{r}|H_{\epsilon\mathrm{MM}})}{P(\mathbf{r}|H_{\mathrm{AMM}})} \tag{10}$$

is the negative log-likelihood ratio [Fig. 2(c)]. In Eq. (9), Γ is the value of the negative log-likelihood which decides the preference of the two models with significance $\alpha = 0.01$. The significance characterizes the type-I error (false-positive) probability [Fig. 2(c)]. It is given by

$$P(\ell(\mathbf{r}) \leqslant \Gamma \mid H_{\mathrm{AMM}} \text{ true}) = \alpha. \tag{11}$$

Taken together, Eq. (9) reads as follows: when $\ell(\mathbf{r}) < \Gamma$, then the AMM hypothesis is not preferred by the data $\mathbf{r}$, in favor the additional complexity offered by the ϵMM hypothesis which significantly explains the variance in $\mathbf{r}$. When $\ell(\mathbf{r}) \geqslant \Gamma$ then the test fails to draw a preference between the two models [Fig. 2(d)], and it cannot be concluded that including non-Markovianity explains the variance in the data. Such an interpretation offered by the negative log-likelihood takes as a basis that the models being compared are nested. Since all ϵ-machines reconstructed via CSSR are done so beginning with a Markovian model and adding more causal states as necessary, the ϵMMs can be considered as containing the AMMs for a set of alternating data. Hence, since the probability of observing a sequence $\mathbf{r}$ given a HMM is efficiently obtained via the forward algorithm [59,69], we were able to compute

the negative log-likelihood ratios with HMM hypotheses in Eq. (10).

Unlike complexity alone, or on and off duration distributions, this statistical framework can determine whether the data cannot be explained by a memoryless process. This only occurs if sufficient data from a non-Markovian process is available (see Appendix B).

D. Non-Markovian memory in brightness state durations

Each time series of state durations [Eq. (2)] consisted of a sequence of state duration times corresponding to each of the seven quantum dots. Since different QD excitation powers induce different physical (and statistical) behavior [8], we consider data from the 100–200 nW range which is expected to behave consistently (see Sec. IV C 3 for further details). Due to the wide range in the sizes of duration times for blinking quantum dots, the amount of unique measurement outcomes $|\mathcal{A}|$ for each time series was on the order of thousands. As mentioned in Eq. (8), large $|\mathcal{A}|$ limits the capacity of CSSR to infer correlations over long past sequence lengths [28]. To combat this, each time series $\mathbf{w}$ was coarse grained over measurement outcomes [Fig. 2(b)]. We performed this by assigning each state duration w^b appearing in a data sequence to a coarse grain symbol, using the rule: $\mathcal{C} : w^b \to c_x^b$ if $w^b \in \{[10^x, 10^{x+1}) : x = 0, 1, 2, \ldots\}$. By using this coarse graining method, each coarse grain symbol c_x^b corresponded to exponentially increasing intervals of brightness state durations:

$$
\begin{aligned}
c_0^b &= [1, 9] \text{ (short)}, \\
c_1^b &= [10, 99] \text{ (medium)}, \\
c_2^b &= [100, 999] \text{ (long)}, \\
c_3^b &= [1000, 99999] \text{ (x-long)},
\end{aligned}
\tag{12}
$$

for $b \in \{\text{On, Off}\}$. Due to the power-law distributed behavior of the state durations $P_b(w^b) \propto w^{-m_b}$, the coarse symbols shared similar behavior, with "short" duration times c_0^b appearing most often, and "long" and "extra-long" durations c_2^b and c_3^b rare occurrences. While we chose to use 10 as the exponent base for the coarse interval sizes, *post hoc* tests showed that the main results of this study were unchanged when we used a moderately larger alphabet. The result of this procedure was a time series of alternating coarse on and off state durations:

$$
\mathbf{c} = \left(c_{x,1}^{\text{On}}, c_{x,2}^{\text{Off}}, \ldots, c_{x,k}^b, \ldots\right),
\tag{13}
$$

shown in Fig. 2(b). The length of our data $\mathbf{c}$ had a range of $N = [428, 1080]$, with a post-compression alphabet size between $|\mathcal{A}| = [6, 8]$. Using Eq. (8) the longest searchable memory length was $L_{\max} = 3$ or 4. The ϵMMs inferred from the QD data in this study are valid as we do not violate this bound.

We constructed the AMMs for each time series of coarse duration times, following the procedure in Sec. III B. The ϵMMs for each time-series data $\mathbf{c}$ were inferred with CSSR over the range $[1, L_{\max}]$, and the non-Markovian memory C_μ recorded. Finally, we performed nLLR tests between the AMM and ϵMM models for each coarse time series, and

recorded the proportion of times the test preferred the non-Markovian model over the AMM.

Table I shows the non-Markovian memory inferred from the data as a function of probed past sequence length $L_{\max}$. When $L_{\max} = 1$, both AMM and ϵMMs had trivial memory for all data, and thus the outcome of the statistical test is not relevant to our investigation. We observe that when $L_{\max} > 1$, the average memory contained in the coarse state durations of blinking QDs is greater than the minimal $C_\mu = 1$ hypothesis, within a spread indicated by the standard deviation $\text{std}(C_\mu)$. Overall, the average non-Markovian memory $\langle C_\mu \rangle$ tended to increase with longer past sequence lengths, which could imply that most quantum dots have nontrivial memory. To confirm this, the hypothesis testing revealed that the AMM is not sufficient to explain the majority of the statistical behavior of the blinking. By assuming the minimal memory via AMMs, the measured data is significantly unlikely to occur. Therefore, we find evidence that most quantum dots have above minimal memory.

Because processes which involve memory can be characterized by it, identifying the underlying drivers or conditions of non-Markovian processes can become more distinct when longer past sequence lengths are taken into account [51]. Similarly, the ability for our methods to rule out memory hypotheses can be improved. Experimentally this is achieved by recording longer data. In our study, we were unable to analyze past sequence lengths greater than $L_{\max} = 4$ due to the limitations in the length of measured data available [Eq. (8)]. However, the methods presented here are scalable for such future work. At long past sequence lengths, the distinction between memory-inducing chemical or physical effects could be made clearer. Diagnosing which of these effects are present in a QD may be achieved by finding theoretical bounds on the amount of memory they are allowed to induce, and comparing those to observations.

The main results presented in Table I only considered laser excitation powers between 100 and 200 nW for consistent physical behavior [70]. Higher excitation powers of 400–600 nW by contrast, are expected to behave differently. When the laser irradiance is higher, the quantum dot is more likely to capture two photons from the same laser pulse. As a result, sufficient energy is available to overcome the activation energy associated with erasing the memory. The expected effect is less memory in the state durations. We do indeed observe this effect for these excitation powers. At 500 and 600 nW, the AMM hypothesis was always preferred over the ϵMM for all probed past sequence lengths (six data sets at 600 nW, 14 data sets at 500 nW). Similarly, less than half (11/23) of the overall data exhibited memory and preferred the ϵMM for 400 nW.

IV. EXPLAINING NON-MARKOVIANITY IN BLINKING

To justify the presence of non-Markovian memory in the blinking quantum dot data, we simulate several hypothetical blinking processes. We determine if they explain the observed non-Marokvianity in Table I. We show physical or chemical process cause a model which has a simple AMM structure, to gain non-Markovian memory. For each memory candidate presented in this section, we simulate its effect on a toy

TABLE I. Results of non-Markovian memory analysis and hypothesis testing of models for blinking quantum dot data as a function of past sequence length $L_{\max}$. For each time series of coarse-grained on and off state durations **c**, the proportion of time series which preferred the ϵMM model over the AMM via the nLLR test is shown. $\langle C_\mu \rangle$ and std(C_μ) represent the mean and standard deviation respectively, of the non-Markovian memory, averaged over each QD at the specified laser power and probed past sequence length $L_{\max}$. See Table II for supplemental time series and memory inference data for all QDs used in this study.

Past sequence length $L_{\max}$	1	2	3	4	Laser power
Proportion preferred ϵMM	0/5	5/5	5/5		
$\langle C_\mu \rangle$	1.000	1.875	1.908		100 nW
std(C_μ)	0.00	0.43	0.40		
Proportion preferred ϵMM	0/5	4/5	4/5	2/2	
$\langle C_\mu \rangle$	1.000	1.573	1.566	2.347	200 nW
std(C_μ)	0.00	0.33	0.33	0.27	

model and show when it strays significantly from the alternating, $C_\mu = 1$ hypothesis. We perform hypothesis testing for each candidate source of memory based on the negative log-likelihood ratio method outlined in Sec. III C. In this section we also rule out systematic data processing treatments we applied as a source of non-Markovianity. All simulations referred to in this section were performed using self-written MATLAB code.

A. Poisson toy model

For a nonblinking quantum dot emitting photons at a constant rate $\lambda_{On} > 0$, the probability of observing n photons arriving in the detector over an observation duration T is Poisson distributed:

$$P_{On}(n) = \frac{(\lambda_{On} T)^n}{n!} e^{-\lambda_{On} T}, \qquad (14)$$

where the n arrival times are distributed uniformly and randomly over T. An ϵ-machine which simulates this process would consist of a single causal state with one self-transition probability corresponding to $P(\text{Arrival}) = \lambda_{On}$ and another corresponding to $P(\text{No arrival}) = 1 - \lambda_{On}$. Having only one causal state, a single Poisson process acting alone in this manner is an example of a memoryless process [71,72]. This corresponds to the starting point of our analysis: that the discrete photon arrivals recorded by the clock are Markovian. For the toy model to simulate a blinking emitter, a Poisson process with a background rate λ_{Off} corresponding to the dark state is added. A chosen master observation duration T is then divided into subintervals of size w. The sizes of the subintervals are power-law distributed,

$$P_b(w^b) = \frac{(w^b)^{-m_b}}{H_{w_*^b}^{(m_b)}} \quad : \quad b \in \{On, Off\}, \qquad (15)$$

where $m_b > 1$ is the power-law constant, and w_*^b is a chosen upper limit on the size of the longest possible subinterval. The constant $H_{w_*^b}^{(m_b)}$ is the generalised harmonic number and is related to the normalization of the distribution. By choosing unique power-law constants m_{On} and m_{Off} in addition to maximum subinterval lengths w_*^{On} and w_*^{Off} for the bright and dark states, we sample the durations $w_1^{On}, w_2^{Off}, \ldots, w_T$ until the master interval is filled. The sequence of alternating subintervals are exactly the state duration time series **w** mentioned in Sec. II B. The arrival times of a blinking quantum dot

are finally simulated by substituting $w_k^b \to T$ into Eq. (14), sampling Poisson arrival times for each brightness duration subinterval in the sequence. Since each Poisson distribution is sampled in an alternating way, the blinking toy model is an example of an AMM described in Sec. III B and shown in Fig. 4, and has $C_\mu = 1$ over duration times as desired. We note that choice of duration distribution in Eq. (15) does not change the memory of an AMM constructed from duration sequences. As long as the duration times between the on and off states are independently distributed we are free to choose whichever distribution we like. We chose the durations to be power-law distributed to match the general behavior of the on and off duration statistics of the experimental data. This procedure resulted in simulated arrival times of a memoryless blinking QD with highly tuneable photoluminescence behavior. The goal of constructing the simulated blinking QD in this way is that we are able to rule out systematic data processing as a source of the observed non-Markovianity. Recalling the data processing outlined in the aforementioned sections, the two main treatments we applied were the binning of the arrival times to intensities [Fig. 1(c)] and the coarse graining of the brightness durations [Fig. 2(b)].
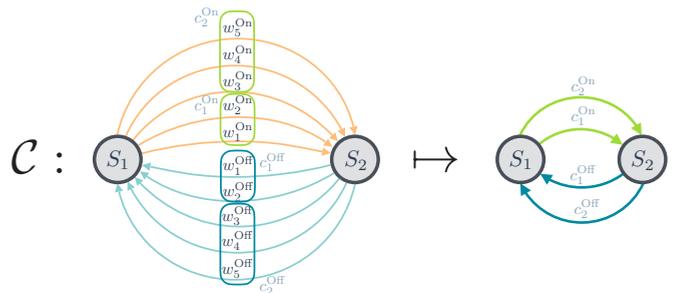

FIG. 4. Example of the effect of a coarse mapping $\mathcal{C}$ on the alternating Markov model. The AMM on the left is at the state duration level. Gold (right pointing) and blue (left pointing) transition probabilities correspond to the "on" and "off" state duration times, respectively. Coarse graining groups together on and off durations w^b into coarse divisions c^b, represented by the boxes, and results in the AMM on the right. The number and distribution of causal states S_i in the AMM before and after coarse graining is the same, and $C_\mu = 1$ in both cases. Hence, as long as the coarse graining only groups together common on and off states, the Markovian property is preserved.

The first treatment is ruled out via the properties of a Poisson process: for an emitter where photons arrive at a constant rate λ_{On}, the probability of observing n photons over *any* time window is Poisson distributed as in Eq. (14). Binning the nanosecond-resolution arrival times into millisecond windows Δt is equivalent to replacing T with Δt in Eq. (14), and thus the statistics of the emitter remain Poisson distributed independent of the size of the binning window. Since Poisson processes are inherently memoryless, this implies they also remain so independently of the choice bin size in counting arrival times.

Figure 4 illustrates that coarse graining the measurement outcomes of an AMM does not increase memory. Starting from the AMM, the possible duration times for the bright state are represented by gold arrows stemming from a single causal state S_1. Similarly, all possible duration times for the dark state are shown as blue arrows stemming from another causal state S_2. Since all the on and off durations each stem from their own respective causal state, this signals that duration times for an intensity level $\{w^b\}$ are statistically independent and identically distributed under some distribution $P_b(w^b)$. Hence, no unique future behaviors are conditional of which specific on and off duration was observed. By applying any coarse graining $\mathcal{C}$ over the duration times in an AMM, the effect is to group transition probabilities stemming from the same state together. The new HMM possesses the same transition structure as the original AMM, where the coarse times in Eq. (12) remain statistically independent, and now identically distributed under some other distribution $P'_{\mathrm{On}}(c^{\mathrm{On}})$ and $P'_{\mathrm{Off}}(c^{\mathrm{Off}})$. In both cases, the number and distribution of the causal states is invariant before and after coarse graining, and hence the non-Markovian memory given by Eq. (7) is not affected. The demonstration of memory invariance under both aforementioned transformations secures the observed memory as stemming from the quantum dot's behavior and/or properties, rather than our methodology.

B. Detector errors

The optics, avalanche photodiode, and clock that collect photon arrival times may have errors which we refer to as "detector errors." To rule out these errors as a cause of memory, we simulated two cases where such errors might be present and examine how a blinking AMM quantum dot may be affected by them. The first is the simple case of background and dark counts, where the photodiode detects a photon arrival when it should not. The second case is the converse *false-off* scenario, where the detector fails to register a photon when it should. This is possible in experiment via a number of ways. First, large portions of photons are not detected due to the limited numerical aperture of the optical system. Some photons are not detected owing to the quantum efficiency of the avalanche diode. Finally, photons may not be detected if they arrive within the dead time of the avalanche diode or the clock, but this is a typically negligible cause.

Assuming that background and false-off detector errors are memoryless and occur independently of the emitting sample [73], their effect on the memory in the state durations is the respective addition and removal of arrivals at a constant rate λ_{err}. For a blinking emitter with photon arrival rates λ_{On} and λ_{Off}, memoryless and independent detector errors will always manifest as a shift in the emitting rates of the brightness states as $\lambda'_{\mathrm{On/Off}} = \lambda_{\mathrm{On/Off}} \pm \lambda_{\mathrm{err}}$. Here the result is another memoryless blinking emitter with Poisson photon arrival rates $\lambda'_{\mathrm{On/Off}}$.

The sum of background photons and dark counts are measured by removing the emitting sample. Adjusting the on and off threshold method [Figs. 1(c), 1(d) and 1(f)] to not classify any measured intensity given under that condition as part of the bright state, characterizes this error. While the false-off error rate in a given experiment involves λ_{On}, the optical aperture, avalanche photodiode efficiency, and efficiency of the optics used, the false-off rates can generally be neglected by choosing QDs that display sufficiently distinguishable brightness states. There are common cases where false-off rates in a detector could be high enough for the bright peak of an intensity histogram to be shifted down to almost background levels. While this type of data is usually discarded, this effect makes the distinction between the brightness states of the QD unclear which in turn has the potential to shift memory away from the AMM $C_\mu = 1$ description. Having obtained data that had a clear separation between the bright and dark states, we find it unrealistic to expect an extreme amount of classification fault our study.

C. Classification

The lack of distinction between brightness states ultimately does influence the amount of non-Markovian memory one can observe. Therefore, we examine which other processes leading to low contrast of brightness states could explain memory found in the blinking patterns. The following subsections deal with potential sources of non-Markovian memory which all have the effect of introducing an ambiguity in brightness state classification during the thresholding step. We simulate four such cases related to plausible physical or chemical processes which may occur in a general blinking QD and examine their influence on memory C_μ, under the change of a key parameter. For each candidate, we inferred an ensemble of ϵMMs from the simulated data and performed nLLR tests to find whether the memory structure deviated significantly from the AMM hypothesis.

1. Low state contrast

Model. Since we classify brightness states takes by choosing a threshold between peaks on a photoluminescence intensity histogram, the simplest way in which brightness state ambiguity may be introduced is via a process where λ_{On} after whatever error, is sufficiently close to λ_{Off}. Overall the effect is realized as an amount of overlap between photoluminescence intensity peaks, which themselves are probability distributions. The photoluminescence intensity histogram peaks are approximately Poisson distributed. This is excluding minor effects which distort the statistics such as excess photoluminescence intensity due to the QD switching brightness states during the binning time window Δt. Assuming exactly Poisson statistics allows us to construct a similarity rating which describes the overlap between the intensity peaks, and quantifies how indistinguishable the bright and dark states are. While many information-geometric quantities exist which all relate

to a distance between probability distributions [74], the Bhattacharyya coefficient (BC) [74,75] is a well-suited starting point for describing overlap. Given two discrete distributions $\mathcal{P}$ and $\mathcal{Q}$ over some domain $\mathcal{X}$, the BC describes the amount of overlap as the intersected probability between them,

$$\text{BC}(\mathcal{P}, \mathcal{Q}) = \sum_{x \in \mathcal{X}} \sqrt{p(x)q(x)}, \quad (16)$$

where $0 \leqslant BC \leqslant 1$, and $BC = 1$ implies the distributions are fully overlapped and hence, identical. For the case of the brightness peaks on a photoluminescence intensity histogram, the full intensity distribution $\mathcal{R}$ is normalized instead of the component bright and dark peaks. For the proportion of intensity events coming from the bright distribution n_{On}, the proportion of events coming from the dark distribution is $1 - n_{\text{On}}$. The full photoluminescence intensity distribution is then described by a convex combination between the bright and dark Poisson curves ($\mathcal{P}$ and $\mathcal{Q}$, respectively),

$$\mathcal{R} \sim n_{\text{On}}\mathcal{P} + (1 - n_{\text{On}})\mathcal{Q}. \quad (17)$$

Referring to the components of the convex mixture as $\mathcal{P}' = n_{\text{On}}\mathcal{P}$ and $\mathcal{Q}' = (1 - n_{\text{On}})\mathcal{Q}$, we define the amount of overlap between them as

$$\text{BCS}(\mathcal{P}', \mathcal{Q}') := 2\sqrt{n_{\text{On}}(1 - n_{\text{On}})}\text{BC}(\mathcal{P}, \mathcal{Q}). \quad (18)$$

Constructing similarity between the bight and dark states in this manner has some nice properties, and interpretation. Effectively it is the regular overlap between the distributions normally expected by the Bhattacharyya coefficient, scaled by how balanced the relative heights the bright and dark Poisson peaks are in the full intensity distribution $\mathcal{R}$. The factor of 2 is included such that $0 \leqslant BCS \leqslant 1$, where $BCS = 1$ implies maximum overlap like for the BC in Eq. (16). Maximum similarity between brightness states is achieved when each state is mixed equally $n_{\text{On}} = 1/2$, and when $\lambda_{\text{On}} = \lambda_{\text{Off}}$. For Poisson-distributed $\mathcal{P} \sim \text{Pois}(\lambda_{\text{On}}\Delta t)$ and $\mathcal{Q} \sim \text{Pois}(\lambda_{\text{Off}}\Delta t)$, a closed-form expression [76] is given by

$$\text{BCS}(\mathcal{P}', \mathcal{Q}') = 2\sqrt{n_{\text{On}}(1 - n_{\text{On}})}\, e^{\left[-\frac{1}{2}(\lambda_{\text{On}}\Delta t - \lambda_{\text{Off}}\Delta t)^2\right]}, \quad (19)$$

which we ultimately use to quantify brightness state overlap in our simulations. The Δt term in Eq. (19) accounts for the freedom of different choices of arrival time binning windows (10ms in our case).

Simulation. The statistical complexity versus λ_{On} was found by simulating photon arrival times of a blinking QD over a master interval of $T = 500\,000$ time steps, and subdividing it into power-law sized subintervals of duration times with constants $m_{\text{On}} = 1.2$ and $m_{\text{Off}} = 1.1$. The relative heights between the bright and dark intensity peaks were modulated by allowing the maximum duration of a single bright state subinterval w_*^{On} to iterate between 2% and 10% of the master interval T. Background rates were fixed to $\lambda_{\text{Off}} = 0.1$ arrivals/time step, with the amount of overlap modulated by varying the bright arrival rate between $0.6 \leqslant \lambda_{\text{On}} \leqslant 2.5$ arrivals/time step. To obtain intensity distributions for the each sequence of simulated arrival times, arrivals were binned over $\Delta t = 10$ time-step windows. Where possible, we defined on and off thresholds using the same methods outlined in Sec. II A for the data. The BCS for each iteration of w_*^{On}

and λ_{On} was calculated. Post thresholding, the simulated data was converted to a sequence of brightness duration times which were then coarse grained using the exponentially scaled bins of base 10 mentioned in Sec. III D. The ϵMMs and complexities C_μ for the simulations were inferred by passing the coarse-grained state duration time series into CSSR at the longest Markov order L_{max} available. The relationship between BCS and C_μ is plotted in Fig. 5(a), in addition to the proportion of simulations with memory structure significantly different to the $C_\mu = 1$ alternating hypothesis as determined by nLLR testing.

Results. Figure 5(a) shows that additional memory above an AMM increases as similarity increases. With a increase in on and off state similarity, the likelihood of observing long brightness durations post-thresholding decreases drastically due to the proliferation of short-length classification errors. Since the area between the on and off intensity peaks corresponds to intensities that have been misclassified, this means long sequences of intensities originally stemming from one brightness state are likely to be interrupted by short misclassifications. For the AMM shown in Fig. 4, the probability of observing a short state duration at time step k was only dependent on which brightness state b a past sequence ended in. With overlap, sequences ending in long durations $\bar{c} = (\ldots, c_{x,k-2}^{b'}, c_{\text{long},k-1}^b)$ have higher probabilities of producing a short duration in the next time step. Since observing a long duration in the last time step $c_{\text{long},k-1}^b$ causes statistically different future behavior compared to seeing any other duration, new causal states and memory are needed to account for this.

We also observe that all simulations significantly prefer the ϵMM over the alternating model for BCS $\geqslant 0.05$. This suggests that very low overlap is required before the simulated quantum dots can no longer be described by the alternating model. It is likely for the data used in this study to have greater overlap than this. Although additional memory simply due to low brightness contrast is a possible effect in our data, it is one that can be tested for in an any general experiment. This is performed by simply observing good separation between photoluminescence intensity peaks, or filtering out quantum dots with low brightness contrast.

2. Time-dependent Poisson rates

Model. Each brightness state of the quantum dot consists of at least two Hamiltonian eigenstates: an excited state and a ground state. It is possible that there are actually many eigenstates with similar, but not quite identical, brightnesses. The thresholding procedure classifies all these eigenstates as the same brightness state. The result allows the emitting rate λ_{On} of the quantum dot to fluctuate randomly in time resulting in an emitting process that is strongly nonstationary [61] Thermal expansion of the optical bench can also cause random changes in the instrument detection efficiency, contributing to this effect.

Simulation. As before, we simulated arrival times by subdividing a master time interval of $T = 500\,000$ time steps into power-law sized subintervals with constants $m_{\text{On}} = 1.2$ and $m_{\text{Off}} = 1.1$, and sampling on and off arrival times within them. We set the single largest possible brightness duration w_*^b for the bright and dark states to 8% and 1% of the master
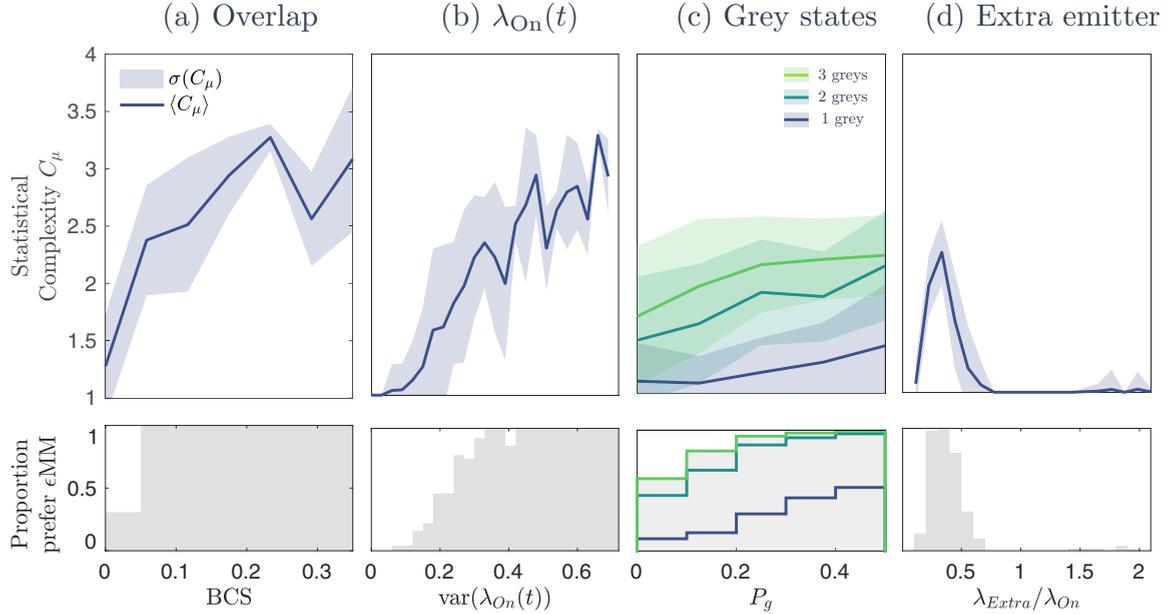
FIG. 5. Memory and model comparison results for each memory candidate. Top rows show mean memory $\langle C_\mu \rangle$ and the standard deviation $\sigma(C_\mu)$. Bottom row shows the proportion of simulations which preferred the ϵMM hypothesis. The scale of the y axes in each row is consistent across subfigures. (a) Memory and model comparison as a function of on and off peak overlap quantified by BCS in Eq. (19). (b) Memory and model comparison as a function of λ_{On} fluctuation strength, quantified by the variance $\mathrm{var}[\lambda_{\mathrm{On}}(t)]$. All simulations with high fluctuation strengths $\mathrm{var}[\lambda_{\mathrm{On}}(t)] > 30$ exhibit memory $C_\mu > 1$. (c) Memory and model comparison vs probability P_g of transitioning to any of the N_g gray states. Each ribbon shows the mean memory and standard deviations for a blinking emitter with 1, 2, and 3 gray states. Each solid line in the model comparison histogram shows the proportion of simulations which could not be significantly explained by the AMM hypothesis relative to the ϵMM, given a number of gray states. (d) Memory and model comparison vs. the brightness of an extra emitter λ_{Extra} relative to a reference $\lambda_{\mathrm{On}} = 2.5$ arrivals/10 time steps.

interval T, respectively. We chose these values of w_*^b following preliminary testing which found them to be suitable for producing a wide range of simulations where on and off states could be classified from the intensity distribution. Arrival times were binned over $\Delta t = 10$ time-step windows. We simulated a fluctuating bright rate by setting λ_{On} to an initial rate of $\lambda_{\mathrm{On}}(0) = 2.5$ arrivals/time step and allowing it to evolve as a random walk between the interval $[0.3, 3.0]$ in steps of 0.1 for each bright state subinterval. The same effect could be achieved by allowing λ_{On} to fluctuate with every time step in the data, not just during bright subintervals. In practice however, locating during which brightness state allows the emitting rate to fluctuates in a blinking QD is nontrivial, hence method suffices as a simplified model. For each iteration of the random walk lower bound, we ran 50 independent simulations. Background rates were fixed to $\lambda_{\mathrm{Off}} = 0.1$. An automatic thresholding script was used to define the on and off states following the classification convention in Sec. II A from the photoluminescence intensity data where possible, after which the simulations were converted to a sequence of brightness durations and coarse-grained using the method in Sec. III D. The ϵMMs and statistical complexities C_μ were inferred from each simulation by sending the coarse-grained time series into CSSR, searching at the largest possible past sequence lengths $L_{\max}$. Electing to keep the random walk step size fixed, we used the lower bound of the random walk as the key parameter which modulated the strength of the fluctuations. This is because decreasing the lower bound

allowed for a higher chance for the variance of the emitting rate fluctuations to increase. By setting the lower bound of the random walk close to the background rate λ_{Off}, the simulated effect could be achieved of the random walk of λ_{On} tending mostly downward.

Results. Using the variance of the emitting rate random walk $\mathrm{var}[\lambda_{\mathrm{On}}(t)]$ as the summary of fluctuation strength, the response of memory C_μ is shown in Fig. 5(b). While an increase in additional memory is possible across most fluctuation strengths, we observe that this is rare when the fluctuations from the initial emitting rate $\lambda_{\mathrm{On}}(0)$ are small. This is evident by the strong grouping of simulations resulting in AMM structures for $\mathrm{var}[\lambda_{\mathrm{On}}(t)] < 0.20$. Above this level, the proportion of simulations with an inferred memory structure significantly different than the alternating $C_\mu = 1$ hypothesis, rises above 50% until it becomes mostly saturated at $\mathrm{var}[\lambda_{\mathrm{On}}(t)] \geqslant 0.35$. Taken together, the results in Fig. 5(b) demonstrate that even when the chemical processes which drive the fluctuations of emitting rates are random, an amount of non-Markovian memory between the lifetimes of brightness levels is incurred. There is no *a priori* reason to expect the random walk of λ_{On} as implemented here is exactly realized in the sense that changes occur only at the beginning of bright states. However, if the emitting rate is allowed to fluctuate freely in time, the extent of the fluctuations may be estimated. This could be done by performing change point analysis (CPA) [77–80] on the intensity data. The number of change points occurring within a band of

high intensity, would give an indication of multiple bright eigenstates.

3. Gray states

Model. Classifying a time series of photoluminescence intensity measurements into bright and dark states is common practice in studies of blinking emitters. While we separated the photoluminescence intensity of the QDs in this study into two brightness states, there may be in general states of intermediate brightness present in quantum dot photoluminescence. The intermediate brightness states, referred to as *gray states* [81]. Gray states occur when a quantum dot undergoes a reversible change in chemical configuration [82], resulting in a lower, but nonzero photoluminescence quantum yield. The number of chemical configurations available is presumably large, but the number of observed quantum yield values is small. Gray states are thought to be present in a variety of systems, and can be identified as excess counts, or a bump between the bright and dark peaks in a photoluminescence intensity histogram (see Fig. 3(b) in Ref. [8]). Gray states are observed for the QDs in this study at high laser powers (300–600 nW). Although the main analysis techniques in our study used lower laser powers which excluded QDs with prominent gray states (Table I), our methodology still generalises to handle the non-Markovianity which may arise from them. Therefore, for this demonstration we set the thresholding procedure to consistently include gray states as part of the dark state. The AMM shown in Fig. 2(f) assumes only two

intensity levels, bright and dark, where a set of coarse-grained duration times is assigned to each level. One might consider adding an extra set of alphabet letters for each additional gray state. However under Eq. (8), larger alphabets limit the length of non-Markovian memory that can be inferred via CSSR. Instead, the durations of the gray states must be embedded within the bright and dark states. This would be likely to involve a memory cost for doing so which we test for here.

Simulation. To simulate a blinking emitter with a number of possible gray states, we followed the same procedure outlined in Sec. IV A, subdividing a master observation window $T = 500\,000$ time steps into power-law-sized brightness durations with constants $m_{\mathrm{On}} = 1.2$ and $m_{\mathrm{Off}} = 1.1$. We allowed the single largest possible brightness duration for the bright and dark states to be fixed to 8% and 1% of the master interval T, respectively. However, instead of sampling photon arrival times alternating between the bright and dark states, we included gray states and followed a simple rule for determining how to transition between intensity levels. Among the QDs in this study that do exhibit gray states, only 1 intermediate level is present [8]. However, other systems such as perovskite quantum dots demonstrate evidence of multiple gray states [83–85], and thus we generalise to that scenario here. Given a specific number of desired gray states N_g, we let P_g be the probability of transitioning to any one gray state from the bright or dark states. For each subinterval brightness duration, the corresponding intensity level was determined via the following $(2 + N_g) \times (2 + N_g)$ transition matrix:

$$
\mathbf{T}_{\text{bright state}} = \begin{array}{c} \\ \text{On} \\ \text{Off} \\ \text{Gray}1 \\ \vdots \\ \text{Gray}N_g \end{array}
\begin{array}{c} \overset{\text{On}}{} \quad \overset{\text{Off}}{} \quad \overset{\text{Gray 1}}{} \quad \overset{\cdots}{} \quad \overset{\text{Gray } N_g}{} \\
\left(\begin{array}{ccccc}
0 & 1 - P_g & \frac{P_g}{N_g} & \cdots & \frac{P_g}{N_g} \\
1 - P_g & 0 & \frac{P_g}{N_g} & \cdots & \frac{P_g}{N_g} \\
\frac{1}{2} & \frac{1}{2} & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\frac{1}{2} & \frac{1}{2} & 0 & \cdots & 0
\end{array} \right)
\end{array}.
\tag{20}
$$

In this example, the simulated quantum dot has a probability of 1/2 to transition to either bright or dark state from any gray state, and for simplicity we deny the possibility for a gray state to transition to another gray state. When P_g is small, the dominant dynamics are alternating between the bright and dark states, with the intermediate gray state transitions acting as noise on the photoluminescence intensity levels. We simulated up to $N_g = 3$ gray states in this example. Each state used power-law sized subintervals with constants $m_g = 1.2$ and allowed the largest state lifetime to be 2% of the master interval T. For the arrival times, we fixed the background and emitting rates to be $\lambda_{\mathrm{Off}} = 0.1$ and $\lambda_{\mathrm{On}} = 2.5$. Photon arrival rates for each gray state were evenly spaced between the interval $[\lambda_{\mathrm{On}}, \lambda_{\mathrm{Off}}]$. Arrival times were binned over $\Delta t = 10$ time-step windows. The probability P_g of entering into any gray state was varied between 0.05 and 0.5. For each iteration we ran 50 independent simulations. Each simulation of intensity data was then sent through a thresholding algorithm, coarse-grained, and passed into CSSR at the longest Markov

order $L_{\max}$ available. The response of C_μ with respect to P_g, in addition to the nLLR testing is shown in Fig. 5(c).

Results. On average, non-Markovian memory increases with the overall probability P_g of entering into a gray state. The increase is shown to be proportional to the number of additional gray states N_g simulated in this example. Both of these effects can be attributed due to the rise in intermediate intensity levels as P_g and N_g increases. It should be noted that while on average memory is increased above $C_\mu = 1$, there are still some instances where nLLR testing fails to draw a preference the ϵMM and the AMM. The lower subplot of Fig. 5 c demonstrates this. For $N_g = 1$ the proportion of times the ϵMM is preferred over the AMM is never greater than 50%, despite high gray transition probability and a mean complexity $\langle C_\mu \rangle > 1$. This effect is reflected in a fraction of the CdSe quantum dots used in this study. At high excitation powers, nearly all quantum dots demonstrated evidence of gray states yet were better described by the AMMs. For multiple gray states however $N_g = 2$ or 3, this is not the case. The proportion

of simulations with memory structure significantly different than the AMM rises and saturates quickly with P_g. Identifying intermediate intensity levels in a quantum dot is nontrivial when the pathways leading to their emissions are rare. In such instances, identifying their presence via photoluminescence intensity histograms alone is difficult and one would turn to more sophisticated methods [8,77,86].

4. Multiple emitters

Model. The final case we consider is when two independent, and blinking QDs are mistaken as a single emitter. This occurs when two quantum dots are chemically bonded together or, by chance, are so close together that an Abbe diffraction-limited microscope cannot distinguish them [87–89]. This effect was simulated by taking two blinking emitters with rates λ_{On} and λ_{Extra}, sampling arrival times independently from each of them, and them superimposing the arrivals. The first emitter was taken as the "reference" with $\lambda_{On} = 2.5$. The second extra emitting rate λ_{Extra} was varied by iterating its value as multiples of λ_{On}, ranging from $0.1\lambda_{On}$ to $2.0\lambda_{On}$. The range of the extra emitter has a dual interpretation. If $\lambda_{Extra} = \lambda_{On}$, then the simulation mimics two quantum dots with the same quantum yield. Otherwise, it mimics two quantum dots with different quantum yields.

Simulation. For each iteration of λ_{Extra}, 25 independent simulations were run. Background rates for were fixed to λ_{Off}. The single longest subinterval brightness duration w_*^b for the reference and extra emitter were set to 8% and 0.6% of the master interval $T = 500\,000$ time steps. For the background, this was set to the usual 1%. Power law constants for the size of the subintervals were $m_{On} = m_{Extra} = 1.2$, and $m_{Off} = 1.1$. After superimposing the reference and extra emitter's arrival times, the data was binned over $\Delta t = 10$ time-step windows to obtain the photoluminescence intensity sequence. These sequences where split into on and off categories and then converted into a sequence of coarse-grained state lifetimes. The memory structure was inferred from the simulations via CSSR, and nLLR tests were performed for each.

Results. Figure 5(d) shows the response of memory C_μ against the ratio $\lambda_{Extra}/\lambda_{On}$, in addition to the nLLR model comparison. Aside from minor statistical fluctuations for $\lambda_{Extra}/\lambda_{On} \geqslant 1.4$, the memory obtained by introducing a second blinking QD is observed to not increase above $C_\mu = 1$ for a large range of relative difference in emitting rates. A high-memory region between $0.2 \leqslant \lambda_{Extra}/\lambda_{On} \leqslant 0.4$ is identified where the majority of simulations prefer the ϵMM over the AMM. At this proportion of relative emitting rates, the extra QD has a rate positioned in between the background λ_{Off} and reference λ_{On} rates. This leads to a similar scenario mentioned in Sec. IV C 1 where the excess of intensity counts between the on and off peaks causes correlations between the state lifetimes, due to classification ambiguity by the threshold procedure. However, the photon antibunching and excitation intensity dependence [8] are distinct from gray state phenomena.

V. CONCLUSION

We have shown that memory is observed in the sequences of on and off state durations by using tools from computational mechanics and hypothesis comparison. By modeling models of memory as hidden Markov models, we have devised a model comparison method and ruled out the AMM model which corresponds to trivial memory in the data. For 90% of the 100 nW and 200 nW QD data used in this study, the statistical test preferred the non-Markovian ϵMM over the AMM when past sequence lengths $L_{max} > 1$ were taken into account (Table I). Preferring the ϵMM over trivial memory indicates the existence of long-range correlations between the amount of time a QD spends in the on or off state. We highlight that the duration histograms used in characterizing QD blinking behavior, as well as correlation functions are unable to reveal non-Markovianity, and we show that observing them to be exponential or power-law distributed gives no information regarding this (Fig. 3).

By pairing our methods with a robust simulation, we were able to simultaneously rule out or associate memory in QD simulations to a number of specific chemical, physical, and experimental effects. Specifically, low state contrast, fluctuating emitting rates λ_{On}, and the presence of gray states were all associated with the presence of memory for simulated data. In relation to the experimental data, gray states were not present in the 100 and 200 nW range, and hence not a cause for the observed memory. Although further experimental analysis outside the scope of this paper is required to determine the relevancy of on and off contrast [Eq. (19)] and fluctuating λ_{On} as cause for memory in our data, we have discussed ways to identify them from QD blinking data in their relevant sections. For example, while the fluctuating emitting rates examined the possibility for similar brightness eigenstates in the QDs, they were also a way of introducing nonstationarity into the simulations. It is known, however, that power-law distributed on and off duration times contribute to the nonstationarity of the intensity dynamics of blinking quantum dots [21,43–45]. Hence, there is room for further exploration of nonstationarity and its relation to memory in future analyses of QDs. We also ruled out time binning, detector errors, and extra identical emitters as effects associated with memory, in both the simulations and experimental data.

It is tempting to assume the aforementioned effects associated with memory may be explained simply due to ambiguity in on and off state classification, observed via an excess in events between on and off peaks [Fig. 1(f)] [90–92]. However, we discussed using the gray state example in Sec. IV C 3 why this is not so. This affirms the notion that the observation of memory in both data and simulation involves more than just any process which results in low on and off state contrast. This counters the argument that memory in blinking QD data can be solely inferred from the intensity histograms, and strengthens the foundation of our analysis.

Although we considered specific examples in Sec. IV C, our methods are general enough to be applied to other chemical systems. One way to test for other intrinsic memory processes would be to perform this analysis on QD data over a range of compositions and measurement techniques. For example, observing memory in systems with blinking rates different to those seen in our data [35], would be an indicator of other chemical processes which involve memory that we have yet to identify. Furthermore, that information could be used to predictably determine the conditions and effects

required for unblinking dots in general. The results for the QDs in this study excited at high laser powers (Sec. III D) present an example where experimental conditions can be associated with non-Markovianity.

Overall, finding memory in QD blinking has allowed us to relate it to specific physical influences that can be experimentally verified. The examples considered in this study are by no means a complete list [82]. Yet we are confident that the rejection of time binning [Fig. 1(c)] and systematic false off or background errors introduced via the detector (Sec. IV B) as sources of memory, ensures the robustness of our methods to be applied to a wide variety of data measured under similar assumptions.

Ultimately what can be predicted is still an open question which will require more collaboration between chemistry and information theory. For quantum dots with memory, each ϵMM has a drastically different graph structure. This difference could be related to differences in quantum dot crystal or surface structure. Our methods do not determine that an ϵMM is a correct model of quantum dot memory. It is possible that substantially larger datasets would reveal a better model, which might be a slight modification of the ϵMM found here, or a model with far more memory. However, the scope of potential questions that may be answered via our methods is large. This is made possible by the scalable design which can readily accept longer data, and requires only that the number and outcomes of measurements are finite. Even if this is not the case, there is a body of theoretical work covering HMMs and memory in such cases [72,93,94], thus making a pathway for our procedure to be extended to more general data. In any case, the observation of memory in QD blinking and its relation to chemical processes in our study, serves as a first step towards discovering a richer and practical understanding of quantum dots. If we can identify a chemical synthesis method that causes quantum dots to "remember" a hidden Markov model that increases their time-averaged quantum yields, then quantum dot lighting will become more efficient.

ACKNOWLEDGMENTS

APPENDIX A: OVERVIEW OF CSSR ALGORITHM

As mentioned in Sec. III B, non-Markovianity in a sequence of observations is related to the dependence of future outcomes to past sequences. The equivalence relation in Eq. (6) formally captures this notion. The CSSR algorithm [28] is able to infer the non-Markovianity in a time series by estimating this equivalence relation. Here we briefly outline how CSSR does this and the relation to two example ϵ-machines.
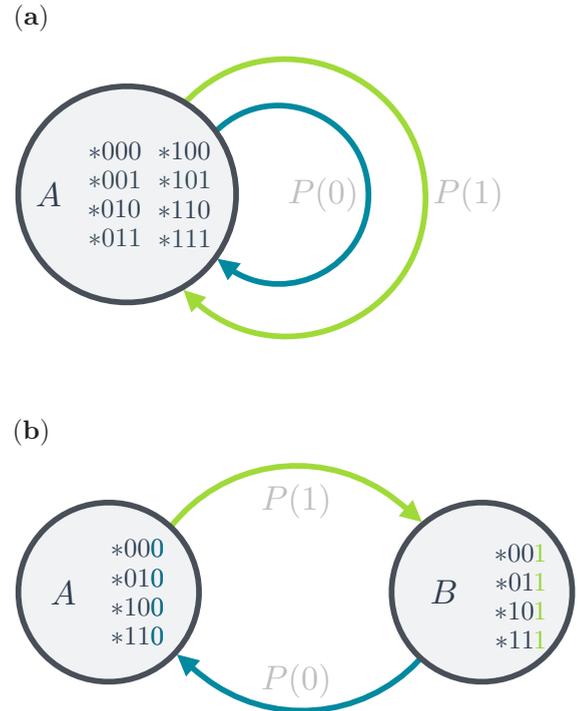

FIG. 6. Example ϵ-machines illustrate the connection between their structure and the equivalence relation in Eq. (6). Both examples consider past sequences up to length $L_{\max} = 3$ of a binary time series where the only possible outcomes in $\mathcal{A}$ are either a 0 or a 1. (a) Example of the ϵ-machine for a memoryless binary time series. Because the future probability of observing a 0 or a 1 are independent of all past sequences of length 3, the single causal state A groups them all together. The $*$ prefix is notation to include any possible sequence containing its suffix. This ϵ-machine is what results from a coin toss—no matter what past sequence has been observed, the probability of flipping a heads $P(1)$ or tails $P(0)$ is unchanged. (b) Example of the ϵ-machine for an alternating binary time series. Here there are two causal states A and B. For this sequence which emits a $010101\dots$ pattern, only the last observation determines the probabilities of future outcomes, even though in this example sequences up to length-3 are considered. Since all length-3 sequences ending in 0 result in equivalent future behavior, they are grouped in causal state A, and vice-versa for those ending in 1.

Recall that the set of all past observations resulting in statistically identical future behavior are grouped together inside the same causal state—this is the statement made by the equivalence relation. A fair coin toss is an example of a process where all past observations lead to the same future behavior. Equivalently, the probability of future outcomes is independent of any past sequence that may be observed. By this notion, there will only be a single causal state that contains all possible past sequences up to some length [Fig. 6(a)], and hence its memory will be zero via Eq. (7). It is clear that a process with memory will be associated with an increase in the number of causal states since some past sequences will require to be *split off* into a separate causal state [see example in Fig. 6(b)]. This is the main concept behind how the CSSR algorithm infers the minimal memory model (the ϵ-machine): it starts by assuming that a time series is memoryless with only
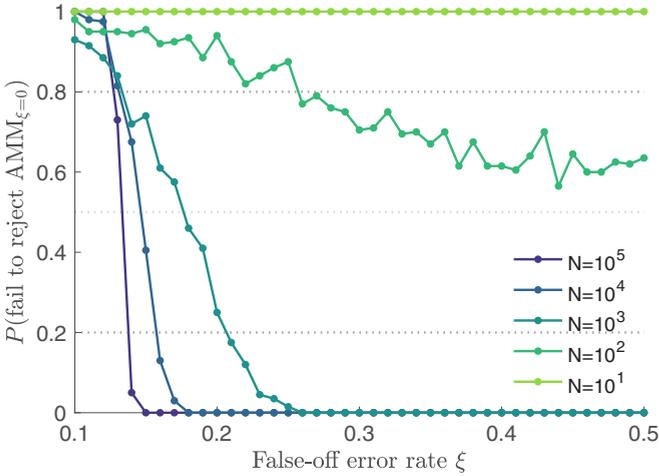
FIG. 7. Power of statistical framework in response to data length and false-off noise in simulations of memoryless blinking QDs. The vertical axis shows the probability of the hypothesis test making a type-II error is $P(\text{fail to reject AMM}_{\xi=0})$. The horizontal axis is the false-off error probability ξ in the simulated data. Hypothesis testing performed using a zero-noise, trivial memory model ($\text{AMM}_{\xi=0}$) as the null and a high-noise, trivial memory model ($\text{AMM}_{\xi=0.5}$) as the alternate hypotheses.

a single causal state and only creates more if Eq. (6) cannot be met.

The algorithm takes in as input a time series $\mathbf{r}$, a library of possible outcomes that can be found in the time series $\mathcal{A}$, the maximum length of past sequences the user wishes to consider $L_{\max}$, and a tolerance parameter $\sigma = 0.001$. The algorithm starts with the memoryless base assumption, and creates a single, empty causal state. Then starting at sequence length $L = 0$ and iterating up to $L = L_{\max}$, the algorithm scans through $\mathbf{r}$ and estimates the probabilities of observing all possible future outcomes $r_k \in \mathcal{A}$ conditional on having observed each past sequence of length L prior: $P(r_k|\bar{\mathbf{r}}_L)$. To test whether the probability of future outcomes based on these past sequences belong in the same or different causal state, the algorithm at each L step compares the aforementioned probabilities to the expected future behavior given any of the past sequences contained in any of the causal states it has so far reconstructed, $P(r_k|\mathcal{S} = S_i)$. This happens by comparing the probability distributions via the *Kolmogorov-Smirnov* (KS) test [95,96]. The hypothesis that $P(r_k|\bar{\mathbf{r}}_L)$ and $P(r_k|\mathcal{S} = S_i)$ are identical is rejected by the KS test at the significance level σ when a distance $\mathcal{D}_{KS}$ is greater than a critical value depending on σ [97]. If no rejection happens for a specific sequence, then it is added as a sequence for the causal state matching its future statistics. If the past sequence results in future behavior significantly different to the sequences currently in all existing causal states, then a new causal state is created to accommodate the future behavior caused by observing the anomalous sequence. Since this splitting of new states happens via the rejection of the KS test, the significance σ represents the type-1 error of mistakenly creating new causal states [98]. Once all past sequences up to $L = L_{\max}$ have been examined and all causal states have been established, the data $\mathbf{r}$ can at each time step be *synchronised* to a specific causal state.

From here CSSR calculates the probabilities from transitioning between causal states. By constructing new causal states only as necessary, the algorithm guarantees a minimal model that describes the non-Markovian behavior of the data (up to a given Markov order), and hence the corresponding ϵ-machine of the process.

APPENDIX B: DATA QUALITY FOR HYPOTHESIS TESTING

The reliability of the hypothesis comparison methods in Sec. III C depends on the length and quality of data sequences available in an experiment. The reliability of a hypothesis test is often characterized by the probability for the test to reject the null when the alternate is true. By exploring this quantity as a function data length and data quality, we devised a scenario using simulated AMMs to provide an estimate on the data conditions needed to use the methods presented in this paper. This is especially relevant in the discussion of power-law distributed processes as we find here, in which many quantities of interest can depend on the observation length. Possible consequences include ageing and ergodicity breaking, and may have implications in the ability of our statistical framework to infer non-Markovianity [9,43].

This was done as follows: A sequence of AMM state durations $\mathbf{w}$ following Eq. (2) was simulated using an AMM [Fig. 2(e)]. Copies of $\mathbf{w}$ were made and false-off errors (Sec. IV B) were introduced at constant probabilities between $\xi = [0.1, 0.5]$. All sequences were compressed using Eq. (12) to form sequences $\mathbf{c}$ following Eq. (13).

We fit two AMMs from the compressed sequences [Fig. 2(f)]: one from the sequence with no error, and another from the sequence with 50% false-off error. To distinguish them, we labeled them by the amount of false off error, e.g., $\text{AMM}_{\xi=0.5}$. We defined $\text{AMM}_{\xi=0}$ as the null hypothesis and $\text{AMM}_{\xi=0.5}$ as the alternate hypothesis. We performed nLLR testing between them using the cloned sequences, truncating them at varying lengths N. Tests were performed at a significance level $\alpha = 0.01$.

The results shown in Fig. 7 demonstrate that an increase in the length of data dramatically improves the ability of the test to reject the null when the data approaches being better described by the alternate model. While for the extreme cases of short data sequences $N = 10^1$ and $N = 10^2$, the hypothesis test struggles to distinguish between the two example models, a length of $N = 10^3$ appears sufficiently accurate. Ultimately the point at $\xi = 0.5$ is when H_1 is true. Since sequences $N \geqslant 10^3$ always reject H_0 at this point, the statistical power of the test is very high in this example.

The point at which the curve goes below 0.5 on the y axis is where the hypothesis test is able to identify perturbations in the data above chance. For increasing N, this occurs at smaller error rates, indicating that the increased sequence lengths increase the sensitivity of the hypothesis testing. For $N = 10^5$ this happens at around $\xi \approx 0.15$ which is assumed to be far above what is generally present from an experiment. This implies that the hypothesis test can be reliably performed without modeling the specific error.

While comparing between the two models mentioned here is a specific example, this demonstration serves as a heuristic

TABLE II. Supplementary time-series information for data in Table I. The total observation time for each dot is the length of the intensity sequence multiplied by 10 ms. Number of blinks is defined as the total number of "On → Off" transitions in the data, not to be confused with the total on and off duration sequence length. Power-law constants for on and off conditions are estimated via a nonlinear fit and shown with 95% confidence intervals. The alphabet size $|\mathcal{A}|$ is the combined number of unique on and off durations in the data, following the compression rule in Eq. (12).

QD	Intensity sequence length (10 ms bins)	Num. blinks	Power-law constant m in Eq. (3) (*On, Off*)	Post-compression alphabet size	Num. states inferred (L_{max} = 1 to 4)	Num. nonzero HMM transitions inferred (L_{max} = 1 to 4)	Laser Power
1	30 025	941	$1.423^{+0.007}_{-0.007}$, $1.650^{+0.005}_{-0.005}$	8	2,7,8	9,22,25	100 nW
4	35 995	631	$1.497^{+0.008}_{-0.009}$, $1.553^{+0.012}_{-0.012}$	8	2,7,7	9,24,25	
5	31 910	784	$1.254^{+0.007}_{-0.007}$, $1.687^{+0.006}_{-0.006}$	8	2,4,4	8,14,14	
6	32 054	989	$1.359^{+0.010}_{-0.011}$, $1.761^{+0.005}_{-0.005}$	7	2,6,6	8,20,20	
7	87 658	843	$1.247^{+0.002}_{-0.002}$, $2.055^{+0.011}_{-0.012}$	8	2,4,5	8,15,18	
1	30 304	880	$1.330^{+0.007}_{-0.008}$, $1.437^{+0.011}_{-0.012}$	7	2,4,4	8,14,14	200 nW
2	46 328	883	$1.002^{+0.043}_{-0.043}$, $1.228^{+0.004}_{-0.004}$	6	2,4,4,15	6,12,12,36	
4	43 615	1080	$1.293^{+0.012}_{-0.012}$, $1.671^{+0.008}_{-0.008}$	6	2,4,4,8	6,12,12,21	
6	20 403	439	$1.242^{+0.018}_{-0.018}$, $1.381^{+0.008}_{-0.009}$	7	2,2,2	8,8,8	
7	36 046	428	$1.363^{+0.010}_{-0.010}$, $1.496^{+0.020}_{-0.021}$	7	2,5,5	9,17,18	

example. In the framework of blinking data in general, we estimate sequence lengths $N \geqslant 10^3$ (i.e., 10^3 blinks) and false-off error rates $\xi < 0.15$ are optimal for our methods. Both of these are achievable in the majority of blinking experiments.

[1] S. W. Koch and L. Bányai, *Semiconductor Quantum Dots* (World Scientific, Singapore, 1993), Vol. 2.

[2] L. Jacak, P. Hawrylak, and A. Wójs, *Quantum Dots* (Springer, Berlin, 1998).

[3] M. Bacon, S. J. Bradley, and T. Nann, Part. Part. Syst. Charact. **31**, 415 (2014).

[4] M. Kouhnavard, S. Ikeda, N. Ludin, N. A. Khairudin, B. Ghaffari, M. Mat-Teridi, M. Ibrahim, S. Sepeai, and K. Sopian, Renew. Sustain. Energy Rev. **37**, 397 (2014).

[5] H. Jun, M. Careem, and A. Arof, Renew. Sustain. Energy Rev. **22**, 148 (2013).

[6] H. Moon, C. Lee, W. Lee, J. Kim, and H. Chae, Adv. Mater. **31**, 1804294 (2019).

[7] J. J. Wierer Jr, N. Tansu, A. J. Fischer, and J. Y. Tsao, Laser Photon. Rev. **10**, 612 (2016).

[8] G. Yuan, D. E. Gómez, N. Kirkwood, K. Boldt, and P. Mulvaney, ACS Nano **12**, 3397 (2018).

[9] M. Kuno, D. P. Fromm, H. F. Hamann, A. Gallagher, and D. J. Nesbitt, J. Chem. Phys. **112**, 3117 (2000).

[10] M. Nirmal, B. O. Dabbousi, M. G. Bawendi, J. Macklin, J. Trautman, T. Harris, and L. E. Brus, Nature (London) **383**, 802 (1996).

[11] S. F. Lee and M. A. Osborne, ChemPhysChem **10**, 2174 (2009).

[12] P. Frantsuzov, M. Kuno, B. Jankó, and R. A. Marcus, Nat. Phys. **4**, 519 (2008).

[13] T. Vosch, J. Hofkens, M. Cotlet, F. Köhn, H. Fujiwara, R. Gronheid, K. Van Der Biest, T. Weil, A. Herrmann, K. Müllen, S. Mukamel, M. Van der Auweraer, and F. C. De Schryver, Angew. Chem. Int. Ed. **40**, 4643 (2001).

[14] A. Efros and D. Nesbitt, Nat. Nanotechnol. **11**, 661 (2016).

[15] M. J. Fernée, P. Tamarat, and B. Lounis, Chem. Soc. Rev. **43**, 1311 (2014).

[16] T. Plakhotnik, M. J. Fernée, B. Littleton, H. Rubinsztein-Dunlop, C. Potzner, and P. Mulvaney, Phys. Rev. Lett. **105**, 167402 (2010).

[17] M. Pelton, D. G. Grier, and P. Guyot-Sionnest, Appl. Phys. Lett. **85**, 819 (2004).

[18] L. Hou, C. Zhao, X. Yuan, J. Zhao, F. Krieg, P. Tamarat, M. V. Kovalenko, C. Guo, and B. Lounis, Nanoscale **12**, 6795 (2020).

[19] J. Houel, Q. T. Doan, T. Cajgfinger, G. Ledoux, D. Amans, A. Aubret, A. Dominjon, S. Ferriol, R. Barbier, M. Nasilowski, E. Lhuillier, B. Dubertret, C. Dujardin, and F. Kulzer, ACS Nano **9**, 886 (2015).

[20] F. D. Stefani, X. Zhong, W. Knoll, M. Han, and M. Kreiter, New J. Phys. **7**, 197 (2005).

[21] J. P. Hoogenboom, J. Hernando, M. F. García-Parajó, and N. F. Van Hulst, J. Phys. Chem. C **112**, 3417 (2008).

[22] N. P. Brawand, M. Vörös, and G. Galli, Nanoscale **7**, 3737 (2015).

[23] D. E. Gómez, J. van Embden, J. Jasieniak, T. A. Smith, and P. Mulvaney, Small **2**, 204 (2006).

[24] S. V. Kilina, P. K. Tamukong, and D. S. Kilin, Acc. Chem. Res. **49**, 2127 (2016).

[25] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry*, 3rd ed. (Elsevier, Amsterdam, 2007).

[26] P. M. Riechers and J. P. Crutchfield, Phys. Rev. Research **3**, 013170 (2021).

[27] J. P. Crutchfield and K. Young, Phys. Rev. Lett. **63**, 105 (1989).

[28] C. R. Shalizi and K. L. Klinkner, in *Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence (UAI'04)*, edited by M. Chickering and J. Y. Halpern (AUAI Press, Arlington, VA, 2004), pp. 504–511.

[29] Y. Ephraim and N. Merhav, IEEE Trans. Inform. Theory **48**, 1518 (2002).

[30] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.106.014127 for plots of normalized event count of on/off durations, brightness histograms, and for intensity versus time recordings of all quantum dots.

[31] K. Boldt, N. Kirkwood, G. A. Beane, and P. Mulvaney, Chem. Mater. **25**, 4731 (2013).

[32] G. Yuan, D. Gómez, N. Kirkwood, and P. Mulvaney, Nano Lett. **18**, 1010 (2018).

[33] E. Lutz, Phys. Rev. Lett. **93**, 190602 (2004).

[34] P. H. Sher, J. M. Smith, P. A. Dalgarno, R. J. Warburton, X. Chen, P. J. Dobson, S. M. Daniels, N. L. Pickett, and P. O'Brien, Appl. Phys. Lett. **92**, 101111 (2008).

[35] S. Ghosh, S. Mandal, S. Mukherjee, C. K. De, T. Samanta, M. Mandal, D. Roy, and P. K. Mandal, J. Phys. Chem. Lett. **12**, 1426 (2021).

[36] S. Stelmakh, K. Skrobas, S. Gierlotka, and B. Palosz, J. Nanopart. Res. **19**, 170 (2017).

[37] S. K. Meena and M. Sulpizi, Langmuir **29**, 14954 (2013).

[38] Z. Fan, A. O. Yalcin, F. D. Tichelaar, H. W. Zandbergen, E. Talgorn, A. J. Houtepen, T. J. Vlugt, and M. A. van Huis, J. Am. Chem. Soc. **135**, 5869 (2013).

[39] T. Graves, R. Gramacy, N. Watkins, and C. Franzke, Entropy **19**, 437 (2017).

[40] L. D. Landau and E. M. Lifshitz, *Statistical Physics*, 3rd ed. (Pergamon Press, Oxford, UK, 1980), Vol. 5.

[41] C. Domb, *The Critical Point: A Historical Introduction to the Modern Theory of Critical Phenomena*, 1st ed. (Taylor & Francis, Oxford, UK, 1996).

[42] H. E. Stanley, Rev. Mod. Phys. **71**, S358 (1999).

[43] X. Brokmann, J.-P. Hermier, G. Messin, P. Desbiolles, J.-P. Bouchaud, and M. Dahan, Phys. Rev. Lett. **90**, 120601 (2003).

[44] G. Margolin and E. Barkai, Phys. Rev. Lett. **94**, 080601 (2005).

[45] F. D. Stefani, J. P. Hoogenboom, and E. Barkai, Phys. Today **62**, 34 (2009).

[46] C. Godrèche and J. Luck, J. Stat. Phys. **104**, 489 (2001).

[47] S. E. Marzen and J. P. Crutchfield, Phys. Lett. A **380**, 1517 (2016).

[48] F. A. Pollock, C. Rodríguez-Rosario, T. Frauenheim, M. Paternostro, and K. Modi, Phys. Rev. A **97**, 012127 (2018).

[49] G. Gennaro, G. Benenti, and G. M. Palma, Phys. Rev. A **79**, 022105 (2009).

[50] C. Stringer, M. Pachitariu, N. Steinmetz, M. Carandini, and K. D. Harris, Nature (London) **571**, 361 (2019).

[51] R. N. Muñoz, A. Leung, A. Zecevik, F. A. Pollock, D. Cohen, B. van Swinderen, N. Tsuchiya, and K. Modi, Phys. Rev. Res. **2**, 023219 (2020).

[52] S. Milz, F. A. Pollock, and K. Modi, Open Syst. Inf. Dyn. **24**, 1740016 (2017).

[53] H.-P. Breuer *et al.*, *The Theory of Open Quantum Systems* (Oxford University Press, Oxford, UK, 2002).

[54] J. Touboul and A. Destexhe, Phys. Rev. E **95**, 012413 (2017).

[55] W. J. Reed and B. D. Hughes, Phys. Rev. E **66**, 067103 (2002).

[56] W. Li, IEEE Trans. Inform. Theory **38**, 1842 (1992).

[57] G. A. Miller, Am. J. Psychol. **70**, 311 (1957).

[58] P. Paradisi, P. Allegrini, F. Barbi, S. Bianco, and P. Grigolini, AIP Conf. Proc. **800**, 92 (2005).

[59] L. R. Rabiner, Proc. IEEE **77**, 257 (1989).

[60] J. L. Doob, *Stochastic Processes* (Wiley, New York, NY, 1953).

[61] W. J. Stewart, *Probability, Markov Chains, Queues, and Simulation* (Princeton University Press, Princeton, NJ, 2009).

[62] P. A. Gagniuc, *Markov Chains: From Theory to Implementation and Experimentation* (Wiley, New York, NY, 2017).

[63] N. F. Travers and J. P. Crutchfield, arXiv:1111.4500.

[64] C. E. Shannon, Bell Syst. Tech. J. **27**, 379 (1948).

[65] R. W. Yeung, *A First Course in Information Theory* (Springer, Berlin, 2012).

[66] CSSR source code repository, accessed May 2021. Retrieved from: https://github.com/stites/CSSR.

[67] CSSR Matlab version, accessed May 2021. Retrieved from: mathworks.com/matlabcentral/fileexchange/33217-causal-state-modeller-toolbox.

[68] K. Marton and P. C. Shields, Ann. Probab. **23**, 960 (1994).

[69] L. Rabiner and B. Juang, IEEE ASSP Mag. **3**, 4 (1986).

[70] N. A. Gibson, B. A. Koscher, A. P. Alivisatos, and S. R. Leone, J. Phys. Chem. C **122**, 12106 (2018).

[71] S. E. Marzen and J. P. Crutchfield, Entropy **17**, 4891 (2015).

[72] S. Marzen and J. P. Crutchfield, J. Stat. Phys. **168**, 109 (2017).

[73] F. Buchinger, A. Kyle, J. Lee, C. Webb, and H. Dautet, Appl. Phys. Lett. **66**, 2367 (1995).

[74] S. Amari, *Information Geometry and Its Applications* (Springer, Berlin, 2016).

[75] A. Bhattacharyya, Sankhya **7**, 401 (1946).

[76] O. Calin and C. Udrişte, *Geometric Modeling in Probability and Statistics* (Springer, Berlin, 2014).

[77] I. M. Palstra and A. F. Koenderink, J. Phys. Chem. C **125**, 12050 (2021).

[78] L. P. Watkins and H. Yang, J. Phys. Chem. B **109**, 617 (2005).

[79] D. L. Ensign and V. S. Pande, J. Phys. Chem. B **113**, 12410 (2009).

[80] D. L. Ensign and V. S. Pande, J. Phys. Chem. B **114**, 280 (2010).

[81] P. Spinicelli, S. Buil, X. Quélin, B. Mahler, B. Dubertret, and J.-P. Hermier, Phys. Rev. Lett. **102**, 136801 (2009).

[82] D. E. Gómez, J. Van Embden, P. Mulvaney, M. J. Fernée, and H. Rubinsztein-Dunlop, ACS Nano **3**, 2281 (2009).

[83] Y.-S. Park, S. Guo, N. S. Makarov, and V. I. Klimov, ACS Nano **9**, 10386 (2015).

[84] B. Li, H. Huang, G. Zhang, C. Yang, W. Guo, R. Chen, C. Qin, Y. Gao, V. P. Biju, A. L. Rogach *et al.*, J. Phys. Chem. Lett. **9**, 6934 (2018).

[85] S. Seth, N. Mondal, S. Patra, and A. Samanta, J. Phys. Chem. Lett. **7**, 266 (2016).

[86] G. Yuan, C. Ritchie, M. Ritter, S. Murphy, D. E. Gómez, and P. Mulvaney, J. Phys. Chem. C **122**, 13407 (2018).

[87] M. J. Rust, M. Bates, and X. Zhuang, Nat. Methods **3**, 793 (2006).

[88] E. Nehme, L. E. Weiss, T. Michaeli, and Y. Shechtman, Optica **5**, 458 (2018).

[89] D. R. Whelan and T. D. Bell, J. Phys. Chem. Lett. **6**, 374 (2015).

[90] I. V. Gopich and A. Szabo, J. Chem. Phys. **124**, 154712 (2006).

[91] I. Gopich and A. Szabo, J. Chem. Phys. **122**, 014707 (2005).

[92] W.-T. Yip, D. Hu, J. Yu, D. A. Vanden Bout, and P. F. Barbara, J. Phys. Chem. A **102**, 7564 (1998).

[93] S. E. Marzen and J. P. Crutchfield, J. Stat. Phys. **169**, 303 (2017).

[94] N. Brodu and J. P. Crutchfield, Chaos **32**, 023103 (2020).

[95] F. J. Massey, J. Am. Stat. Assoc. **46**, 68 (1951).

[96] M. Hollander, D. A. Wolfe, and E. Chicken, *Nonparametric Statistical Methods*, 3rd ed. (Wiley, New York, NY, 2013).

[97] L. H. Miller, J. Am. Stat. Assoc. **51**, 111 (1956).

[98] Preliminary testing demonstrated the output of CSSR to be robust under small changes in σ. Our analysis used the default value $\sigma = 0.001$.