 doi:10.20944/preprints201905.0327.v1

Article

Systematic assessment of freely-diffusing single-molecule fluorescence detection using Brownian motion simulations

Dolev Hagai [1] and Eitan Lerner [1,*]

[1] Department of Biological Chemistry, The Alexander Silberman Institute of Life Sciences, Faculty of Mathematics & Science, The Edmond J. Safra Campus, The Hebrew University of Jerusalem, Jerusalem, Israel; Eitan.Lerner@mail.huji.ac.il

* Correspondence: Eitan.Lerner@mail.huji.ac.il; Tel.: +972-2-658-5457

Abstract: Single-molecule fluorescence detection (SMFD) experiments are useful in distinguishing between sub-populations of molecular species in measurements of heterogeneous samples. One of the experimental platforms for SMFD is based on a confocal microscope setup, where molecules in the solution randomly traverse an effective detection volume, formed by a tightly focused laser beam. The non-uniformity of the excitation profile and the random nature of Brownian motion, produce fluctuating fluorescence signals. For these signals to be distinguished from the background, single-molecule fluorescence burst analysis is frequently used. Yet, the relation between the results of burst analyses and the underlying spatial information of the diffusing molecules is still obscure and requires systematic assessment. In this work we performed three-dimensional Brownian motion simulations of SMFD, and tested the positions from which the molecules emitted photons that were detected and passed the burst analysis criteria for different values of burst analysis parameters. The results of this work verify which of the burst analysis parameters and experimental conditions influence both the position of molecules in space when fluorescence is detected and taken into account, and whether these bursts of photons arise purely from single molecules, or not entirely.

Keywords: single-molecule; fluorescence; burst; photon rate; effective detection volume; point-spread function; Brownian; diffusion; simulation; threshold.

1. Introduction

The outstanding capability to detect fluorescence from single molecules one at a time allowed distinguishing between different molecular species based on many parameters derived from these experiments[1]. In the field of single-molecule fluorescence detection (SMFD) there are two main experimental platforms: i) fluorescence imaging of single molecules immobilized to the surface of a coverslip; and ii) detection of fluorescence bursts from freely-diffusing single molecules through a small effective detection volume. The advantage of immobilized SMFD over its freely-diffusing counterpart is that the fluorescence from each individual molecule can be recorded for long periods of time. Still, the main limitation of immobilized SMFD is the inherent requirement to immobilize the molecule in the first place. Within this perspective, the common immobilization is through chemical conjugation, which always raises criticism about the possible perturbation this procedure might cause to the biomolecule under measurement. There are other techniques that allow limiting a biomolecule to a small volume, such as the anti-brownian electrokinetic (ABEL) trap[2], or confinement by encapsulation inside vesicles[3,4] or liposomes[5]. Nevertheless, each such method presents its own possible limitations (the electric field used in the ABEL trap to center back the molecules might affect them; it might be hard to exchange chemical conditions if the molecules are encapsulated in liposomes). For the abovementioned reasons, some researchers choose the confocal-based SMFD of freely diffusing molecules. Additionally, the amount of different single molecules probed in this version of SMFD can be potentially higher than in the immobilized SMFD, which

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 May 2019 doi:10.20944/preprints201905.0327.v1

makes it better in applications that require a higher throughput. Still, the nonuniformity of the effective excitation at different positions in space and the random Brownian motion of molecules into these different regimes in space, make it hard to distinguish different signals from single molecules from each other and from the experimental background.

A confocal-based SMFD measurement of freely diffusing molecules involves laser excitation tightly focused through a high numerical aperture objective lens (usually water immersion), where the focus is brought deep (a few tens of μm) inside the sample solution. Then, if the overall laser power is low, effective excitation will occur mostly when molecules traversed through the laser focus. Lenses and a narrow pinhole (diameter of a few tens of μm) focus fluorescence photons, collected by the objective lens, so that only the maximum of focused fluorescence crosses it, while its periphery is rejected. Then, fluorescence is re-collimated by additional lenses, spectrally selected (with proper filters) and focused, so that light reaches the active area of single-photon avalanche diodes (SPADs; the detectors). Confocal-based SMFD measurement of freely diffusing molecules is carried out by a setup as described above, based on the premise that most of the time no molecule traverses the effective detection volume, and when some do, they are mostly single ones. To promise this condition is kept, these measurements are performed in concentrations of the fluorescently labeled molecules (tens of pM, or lower). The measured trace is comprised of a long list of detection times, relative to the moment the recording started, as well as detection tagging, usually used in SMFD applications, in which more than one SPAD is used (such as in single-molecule Förster resonance energy transfer; smFRET). In summary, most of the measured trace is background, since most of the time no molecule crosses the effective detection volume, and there are short intervals of time, in which fluorescence from molecules are detected. Therefore, an inherent requirement of SMFD of freely-diffusing molecules is that the background will be low and the fluorescence signal will be high. To keep the signal high, one needs to choose fluorophores with a high absorption coefficient at the wavelength of excitation and a high fluorescence quantum yield. This set of requirements is also known as fluorophores with higher *molecular brightness*. To explain how to keep the background low, we first have to understand what the sources of background are. The dark counts of SPADs comprise a part of the measured background, however other sources could be due to molecules excited out of the focus, as well as light scattering that was not optically rejected efficiently enough. If both the background and the signal comprise of fluorescence signals, what differs in-focus from out-of-focus fluorescence? The answer could be hidden in photon detection rates. If the time interval of all pairs of consecutive detection moments will be collected into a histogram, one will identify it includes two time-dependent processes – a slow Poisson process and a fast Poisson (and sometimes super-Poisson) process. If so, identifying fluorescence from molecules that traversed the focus region of space (the effective detection volume) is a matter of identifying signals with instantaneous photon detection rates high enough relative to the mean rate of the slow Poisson process of the background.

This procedure is precisely the one used in burst identification in SMFD of freely-diffusing molecules. The single-molecule burst analysis procedure of SMFD has been discussed in many previous works[6–11]. The procedure includes a window of m consecutive photon detections, for which we calculate the time, t, from the first detection time (No. 1) to the last one (No. m) [6,12], or the rate, f, by dividing m by this time interval[11]. In the analysis, this window slides one photon detection event at a time. These calculations are used for the definition of the signal as photon bursts with instantaneous photon detection rates higher than a given threshold. Doing so, the instantaneous time, t, for the m consecutive detection events should be smaller than some arbitrary time T[6,12]. Using the m-rate notation, the rate, f, should be higher than some arbitrary rate[11]. To make the threshold choices a bit less arbitrary, one can choose a minimal rate threshold, F times higher than the background rate. This way, at least, the signal is defined relative to the background objectively. Still, the choice of the value of this F parameter is arbitrary.

After identification of photon bursts, each burst is defined by the total amount of photons in it (the burst size), the time interval from the first to the last photon in the burst (the burst width), the ratio of the burst size and width (the mean photon brightness) and the maximal instantaneous photon rate (also equivalent to the molecular brightness). After burst identification, it is customary to filter photon bursts according to some of these criteria or ratios of them, using, again, some arbitrarily

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 May 2019 doi:10.20944/preprints201905.0327.v1

chosen values. Many choose the burst size as a burst selection criterion, where the larger the burst size is, the more trustworthy it is in calculations using its size, such as in the calculations of the mean FRET efficiency using burst sizes recorded in two SPADs. However, since in SMFD experimental, conclusions are drawn from burst-dependent histograms (e.g. FRET histograms, burst width histograms), many bursts are required to pass the burst selection criterion, but the higher the minimal burst size threshold is, the lower the amount of bursts that will be selected.

Overall there is a price for arbitrary choice of parameter thresholds in this process, and it is important to understand what is the meaning of choosing different threshold values at the level of the chosen molecules and their positions in space, when they emitted a photon that was selected as part of a burst. In this work we used PyBroMo (python Brownian motion) simulations (https://github.com/tritemio/PyBroMo; was utilized in previous works[13–16]) to simulate SMFD of freely-diffusing molecules. Doing so, we record both the molecular positions in the simulation and the photon detection times. Then, after employing the burst analysis procedure, as we usually do in experiments, we assess the positions of single molecules, at the moments when each photon of each burst was detected. By doing so, it was possible to define the spatial maps of positions as a function of the burst analysis parameters. We tested different *m*-photon windows, different *F* factors, different burst size and different burst width minimal thresholds. Additionally, we repeated the assessment for different simulations results with different diffusion coefficients at a given molecular concentration, as well as different concentrations at a given diffusion coefficient. We show that in order to identify bursts of molecules that traversed a well-defined narrow region of the effective detection volume, where the bursts are constructed mostly of photons of single molecules, the best practice would be to use large values of photon detection rate thresholds, *F* and modest values of the burst size threshold. We also show that measuring lower concentration of molecules further improves the single-molecule burst identification. Finally, we show that caution has to be taken with the use of large values of *F*, when molecules with low diffusion coefficients are being measured.

2. Results

We used the Python Brownian motion (PyBroMo) code to perform 3D diffusion simulations, where we record the x, y & z positions of each diffusing molecule at each moment and advance the molecular positions in intervals of 200 ns. The diffusion simulation times were 60 seconds. Table 1 summarizes the different simulation conditions we tested.

Table 1. Different 3D diffusion simulation conditions that were tested.

Box[1] x (μm)	Box y (μm)	Box z (μm)	Box vol.[2] (fL)	No. of molecules	Conc.[3] (pM)	D[4] (μm²/s)
5.54	5.54	13.85	425	15	62	90.0
6.60	6.60	19.70	858	15	31	90.0
8.30	8.30	24.90	1715	15	15	90.0
5.54	5.54	13.85	425	15	62	22.5
5.54	5.54	13.85	425	15	62	5.6

[1] The molecules diffuse inside a box having a length x, depth y & height z.

[2] The box volume is calculated from the box dimensions x, y & z

[3] The concentration is derived from the No. of molecules divided by Avogadro's number and then divided by the volume (in L units)

[4] The diffusion coefficient.

To model the effective detection volume we considered the model of a realistic point-spread function (PSF) of a typical 60x water immersion objective with a numerical aperture of 1.2, with a sample mounted on top a 150 μm coverglass and with sample excitation at a wavelength of 532 nm (see Figure S1). We modeled the PSF using a vectorial electromagnetic simulation, PSFLab[17]. This model includes effects of refractive index mismatch as well as mismatch between objective lens correction and coverglass thickness.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 May 2019 doi:10.20944/preprints201905.0327.v1

Then, in the next step, the instantaneous emission rates of each molecule at each moment in the Brownian motion simulation were calculated by evaluating the intensity of the PSF at different points in space, which dictate the probability for excitation, emission and detection. After the diffusion simulation and the calculation of the instantaneous emission rates, photons were generated from a Poisson process using the instantaneous emission rates. Each detected photon was assigned the molecule from which it originated. Afterwards, Poisson background timestamps were also added. In the end of the simulation, a photon HDF5[13] file was constructed, including all photon timestamp assignments, as if it was a file containing experimental data. Therefore, next, each simulated photon HDF5 file was analyzed for identifying single molecule photon bursts, using the FRETbursts single-molecule photon detection analysis suite[18].

The burst analysis parameters that were tested and their values were: i) the number of consecutive photons, m, in the burst search sliding window (values 5, 10, 15 & 20, at a constant photon rate threshold, F=6), ii) the photon rate threshold, F, defined as the minimal multiplier of the background rate that can be considered a signal (values 3, 6, 11, 16 & 21, at a constant m=10 consecutive photons), iii) minimal burst size thresholds (values 10, 20, 40 & 80, for constant m=10 and F=6; burst size threshold value of 10 is already included for m=10 & F=6, since by definition the identified bursts include at least m=10 consecutive photons) and iv) minimal burst width thresholds (values 0.0, 0.5 & 1.0 ms, for constant m=10 and F=6; burst width threshold value of 0.0 ms is already included for m=10 & F=6, since by definition the identified bursts include all bursts, with all burst widths).

2.1. Molecular position dispersion

Because the simulations allowed recording the position of each molecule at any instance of the diffusion simulation, including instances in which the molecule yielded a detected photon, we plotted the positions of the molecules when they emitted a photon that was detected and was part of a single molecule burst using the burst analysis parameters shown above. This produced a 3D scatter plot of positions, where we chose to show the xy, xz & yz scatter plot projections for z=0, y=0 & x=0, respectively. Then, we have overlaid the contour line of the 90% of the PSF in these 2D projections, for reference (Figures 1 & 2). These results are for the simulation of 15 molecules in 425 fL rectangular box (yielding a concentration of 62 pM), where the diffusion coefficient of the molecules was 90 $\mu m^2/s$.

Figure 1 shows the positions of the diffusing molecules at the moment they emitted a photon, after they have been selected as belonging to bursts, by minimal burst analysis criteria using a sliding window of m=5 consecutive photons, and identifying bursts where the instantaneous photon rate was at least F=6 times higher than the background rate, and with no additional burst filtration.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 May 2019 doi:10.20944/preprints201905.0327.v1

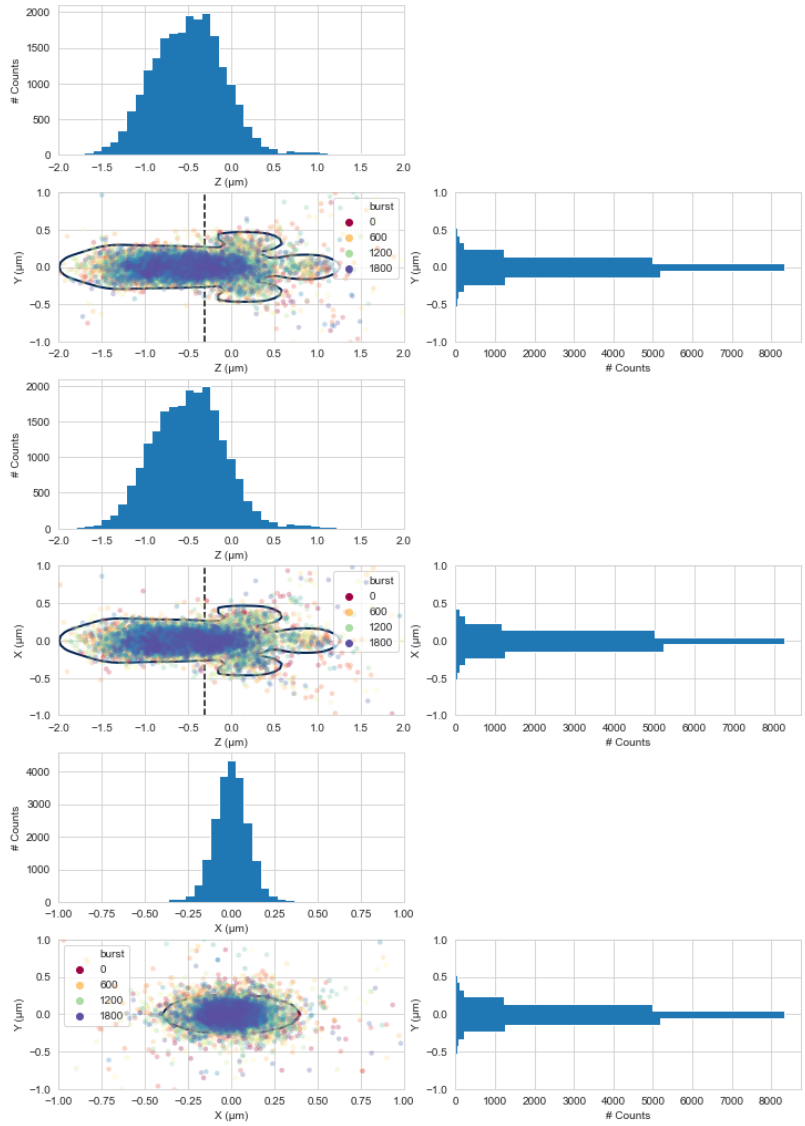


Figure 1. The positions of diffusing molecules when they emitted photons that were detected and selected by the burst analysis, with burst analysis parameters m=5 & F=6. In the top, central & bottom panels we show the 2D projections at the yz, xz & xy planes when x=0, y=0 & z=0, respectively. Each dot in the scatter plots is an emitted photon. These results are for the simulation of 15 molecules in 425 fL rectangular box (yielding a concentration of 62 pM), where the diffusion coefficient of the molecules was 90 μm²/s. The colors of the points correspond to the burst number out of the overall number of bursts. In each panel, the 1D projections are also shown as histograms.

Figure 2, on the other hand, shows the positions of the diffusing molecules at the moment they emitted a photon, after they have been selected as belonging to bursts, by stringent burst analysis criteria using a sliding window of m=10 consecutive photons, and identifying bursts where the instantaneous photon rate was F=6 times higher than the background rate, and after selection of only bursts that had a size of more than 40 photons.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 May 2019 doi:10.20944/preprints201905.0327.v1

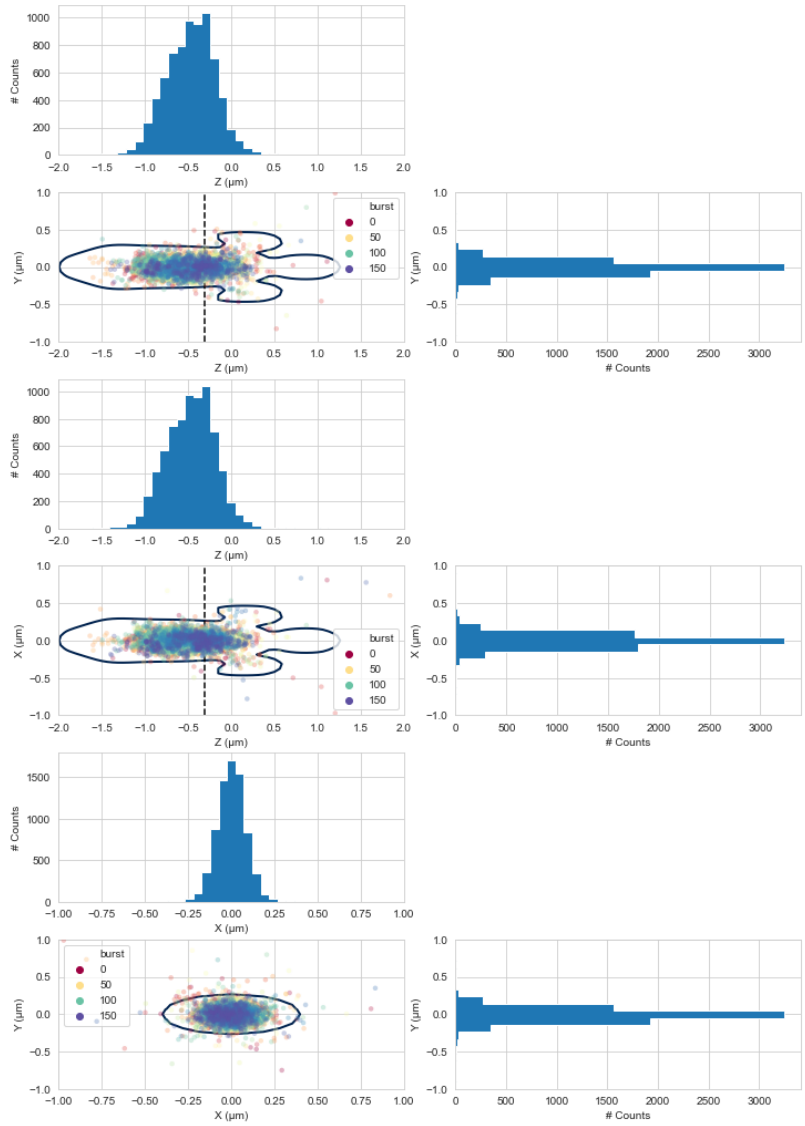


Figure 2. The positions of diffusing molecules when they emitted photons that were detected and selected by the burst analysis, with burst analysis parameters m=10, F=6 & a minimal burst size threshold of 40. In the top, central & bottom panels we show the 2D projections at the yz, xz & xy planes when x=0, y=0 & z=0, respectively. Each dot in the scatter plots is an emitted photon. These results are for the simulation of 15 molecules in 425 fL rectangular box (yielding a concentration of 62 pM), where the diffusion coefficient of the molecules was 90 μm²/s. The colors of the points correspond to the burst number out of the overall number of bursts. In each panel, the 1D projections are shown in histograms.

In both cases, it is clear that the positions of the molecules are well within the PSF with a tendency towards its center. Therefore, the burst analysis procedure achieves its primary goal – identifying signals from molecules, when they traversed the effective detection volume. However, while in the case of minimal burst analysis criteria the position dispersion (i.e. the spread or the dispersion of the positions) was quite wide, in the case of stringent burst analysis criteria, the position dispersion became smaller. Note that the fraction of events that were detected outside the PSF region using minimal burst analysis criteria was much larger than that when using stringent burst analysis criteria. How should one go about deciding which criteria are considered stringent and which are not? How might one reduce arbitrary value choices? To answer these questions, it is important to assess which of the burst analysis parameters are the parameters that highly affect the reduction in the position dispersion around the center of the PSF.

Figure 3 shows the 1D histograms of the molecular positions as a function of the z coordinate (Figure 3, left panels) and the x coordinate (Figure 3, right panels), for the four different tests of the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 May 2019 doi:10.20944/preprints201905.0327.v1

burst search parameter values explained above (from top to bottom, different m values with a constant F, different F values with a constant m, different minimal burst size and different minimal burst width thresholds with constant m & F).

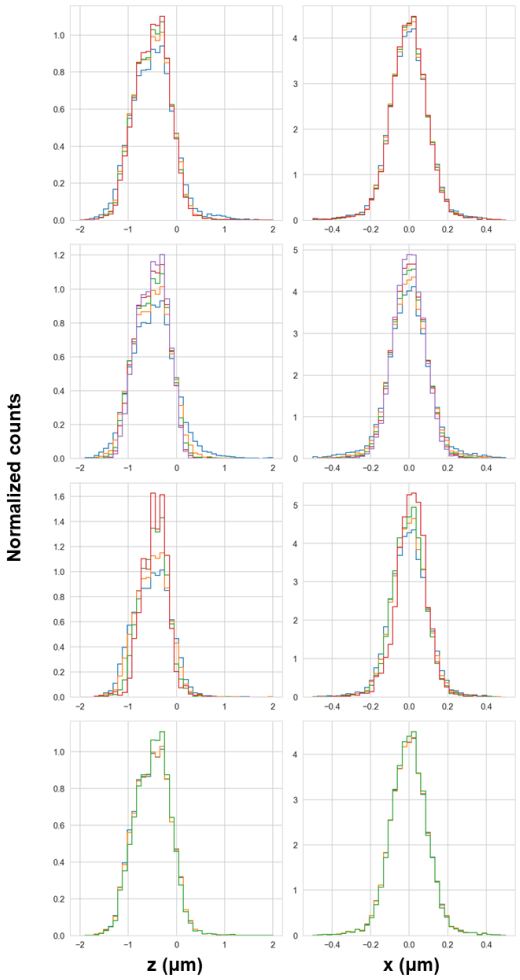


Figure 3. histograms of the 1D projections shown in Figures 1 & 2 for the z & x coordinates (left & right panels, respectively). From top to bottom, we assessed these histograms as a function of a sliding window of m consecutive photons (m=5, 20, 15 & 20 in blue, orange, green & red, respectively), using a constant instantaneous photon rate threshold of F=6; as a function of the instantaneous photon rate threshold F, (F=3, 6 ,11 ,16 & 21 in blue, orange, green, red & magenta, respectively), using a sliding window of constant m=10 consecutive photons; as a function of the minimal burst size threshold (10, 20, 40 & 80 in blue, orange, green & red, respectively); and as a function of the minimal burst width threshold (0.0, 0.5 & 1.0 ms in blue, orange & green, respectivey), for a constant m=10 & F=6. These results are for the simulation of 15 molecules in 425 fL rectangular box (yielding a concentration of 62 pM), where the diffusion coefficient of the molecules was 90 $\mu m^2/s$. The colors of the points correspond to the burst number out of the overall number of bursts.

Qualitatively, increasing the different burst analysis values makes the position dispersion smaller, however to different degrees. It is clear, for instance, that increasing the instantaneous photon rate threshold, F, and the minimal burst size threshold have the largest effect on decreasing the position dispersion. It is also clear that the minimal burst width threshold does not have a significant effect on the molecules' position dispersion.

To quantitatively assess the effect of choice of burst search parameter values on the position dispersion, we quantified the position dispersion in each dimension by calculating the standard deviations of the x, y & z positions. Figure 4 reports the results of quantifying the molecules' position dispersion as a function of the burst analysis parameter values assessed in the histograms of Figure 3.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 May 2019 doi:10.20944/preprints201905.0327.v1

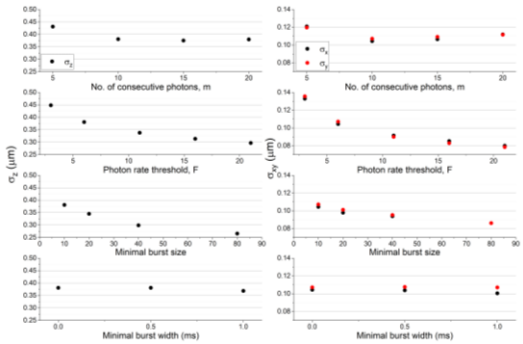


Figure 4. Quantification of the position dispersion in x, y & z using different burst analysis parameter values. The quantification was achieved via calculation of the standard deviation of the molecule positions in x , y (right) & z (left). From top to bottom, we quantified the molecules' positional narrowing as a function of a sliding window of *m* consecutive photons, using a constant photon rate threshold of *F*=6; as a function of the photon rate threshold *F*; as a function of the minimal burst size, , for burst search results using *m*=10 & *F*=6; and as a function of the minimal burst width, , for burst search results using *m*=10 & *F*=6. These results are for the simulation of 15 molecules in 425 fL rectangular box (yielding a concentration of 62 pM), where the diffusion coefficient of the molecules was 90 μm²/s. The colors of the points correspond to the burst number out of the overall number of bursts.

One can see that among all factors assessed here, increasing the photon rate and minimal burst size thresholds decreases the position dispersion the most. This means, that the burst analysis retrieves bursts of molecules that were acquired while they were traversing the effective detection volume, and the higher the values of these thresholds were, the closer their position was to the center of the effective detection volume.

2.2. Pure and impure single-molecule bursts

The approach used here is built on the premise that the photons in each single molecule burst are emitted from a single molecule, and hence can be considered *pure single-molecule bursts*. However, it is certainly possible that there will be a fraction of single molecules that will include photons that arise mostly from a single molecule, whereas a few photons from another molecule that crossed the effective detection volume at the same time. There is also a possibility that other molecules traversing the laser beam region, out of the focus, will emit photons, with a low rate. With the abovementioned two possible scenarios, one can imagine a molecule traversing close to the center of the effective detection volume, hence emitting photons at a high rate, and another molecule out of focus, or at the rim of the effective detection volume, also emitting photons at the same moments, but at a lower rate. The overall photon emission rate will be the combined photon rate, which is why this burst might be selected by the analysis. We term such bursts, *impure single-molecule bursts*.

According to the possible explanations given for the formation of impure single-molecule bursts, three factors may influence the occurrence of such bursts and their level of impurity: 1) the choice of the burst search parameter values; 2) the concentration; and 3) the diffusion coefficient. While the first factor is intrinsically controlled in the data analysis, the second & third factors are extrinsic parameters that can be controlled in the experiments.

In Figures 1-4 we have shown that by controlling the burst analysis parameter values we modify the molecular positions that are selected as photons in bursts. If an impure single-molecule burst becomes impure due to an additional molecule that has a low photon rate because it is outside the effective detection volume, or on its rim, reduction of the position dispersion will lead to reduction in the occurrence of such bursts, and to a reduction in the level of impurity.

To assess this suggestion, we calculated the level of impurity in a burst as the complement of the ratio of the amount of photons arising from the most frequent molecule and the overall amount of photons (as in Eq. 1):

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 May 2019 doi:10.20944/preprints201905.0327.v1

$$f = 1 - (\text{photons from most frequent molecule})/(\text{all photons}), \quad (1)$$

Background timestamps were discarded from the calculation of amount of photons. Figure 5 shows the histograms of bursts as a function of the level of impurity, expressed as the fraction of photons from other molecules (as in Eq. 1).

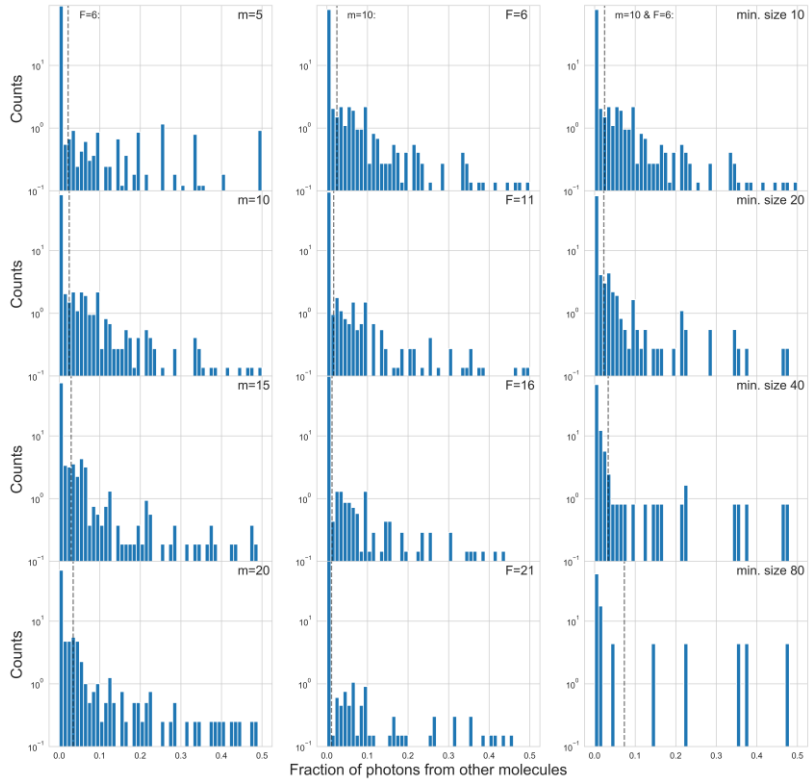


Figure 5. Quantification of the level of impurity of single-molecule bursts using different burst analysis parameter values. Each panel shows a histogram of all the bursts' level of impurity, calculated as the fraction of photons arising from molecules other than the main one. The burst impurity histograms are of the simulation results after burst search analysis using a constant photon rate threshold F=6 and varying value of m, for a sliding window of m consecutive photons (left), constant m=10 and varying F values(center) and constant m=10, F=6 and varying burst size threshold values (right). The dashed vertical grey lines indicate the mean impurity value for all bursts. These results are for the simulation of 15 molecules in 425 fL rectangular box (yielding a concentration of 62 pM), where the diffusion coefficient of the molecules was 90 $\mu m^2/s$. The colors of the points correspond to the burst number out of the overall number of bursts.

Qualitatively, one can make the following inferences: 1) increasing the value of m, for a sliding window of m consecutive photons, used in the burst search process, increases the amount of impure bursts as well as increase the burst impurity overall; 2) increasing the photon rate threshold, used in the burst search process, decreases the impure bursts and the leftover impure bursts have a lower level of impurity; 3) increasing the burst size threshold leaves us with more impure bursts, but with a lower level of impurity. Overall, increasing the bursts with higher sizes include more photons arising from molecules other than the main one in the burst. This is because, the higher the size is, the higher the probability that one of the photons arose from a different molecule, owing to its low photon rate. Naturally, increasing the minimal photon rate (F) allows better rejection of such instances.

Two main inferences can be made from burst impurity histograms as the ones in the panels of figure 5: 1) the fraction of bursts that are pure (impurity level of 0) and the fraction that are impure (impurity level higher than 0); and 2) the mean level of impurity, as the mean of all the impurity

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 May 2019 doi:10.20944/preprints201905.0327.v1

values of all the bursts (Figure 5, dashed vertical grey lines). Figure 6 summarizes the values of these parameters for different burst search parameters.

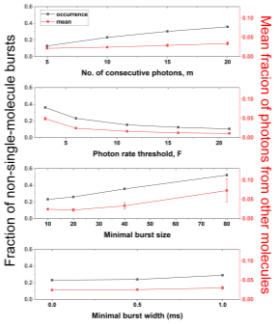


Figure 6. The occurrence and level of impure single-molecule bursts using different burst analysis parameter values. The relative occurrence of impure bursts was calculated as the fraction of bursts with an impurity level larger than 0, as the fraction of non-single-molecule bursts. The man level of impurity was calculated as the mean of all burst impurity levels, as the mean fraction of photons from other molecules (other than the main molecule in a burst). These results are for the simulation of 15 molecules in 425 fL rectangular box (yielding a concentration of 62 pM), where the diffusion coefficient of the molecules was 90 $\mu m^2/s$. The colors of the points correspond to the burst number out of the overall number of bursts.

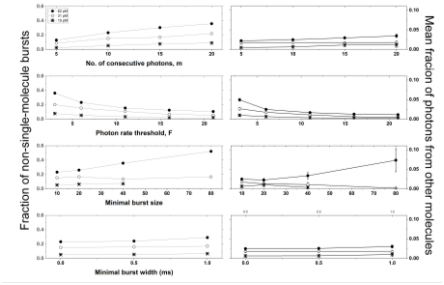
One can clearly see the description qualitatively portrayed above, is quantitatively justifiable. Therefore, although increasing the values of both the photon rate threshold, *F*, and the burst size threshold, leads to the desired decrease in the molecular position dispersion, mainly increasing the value of *F* and moderately increasing the value of the burst size threshold keeps the occurrence of impure bursts low.

2.3. Experimental control over the molecular position dispersion and on the occurrence and level of impure single-molecule bursts

As mentioned above, in addition to controlling the results by using different burst analysis parameter values, they can also be controlled experimentally by decreasing the concentration or decreasing the diffusion coefficient. We assessed the effect of both the experimental and analytical control over the results by performing the simulations either in decreasing concentrations for a constant diffusion coefficient or in decreasing values of diffusion coefficient for a constant concentration.

2.3.1. Decreasing concentrations further decreases the molecular position dispersion and the occurrence of impure single-molecule bursts

Figure 7 quantitatively proves what was conceptually obvious: the lower the concentration of the measured molecules is, the lower the occurrence of impure bursts will be. Additionally, for the leftover impure bursts that passed the burst selection, their level of impurity will be smaller, in comparison to a measurement in higher concentrations.



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 May 2019 doi:10.20944/preprints201905.0327.v1

Figure 7. The occurrence and level of impure single-molecule bursts decrease when decreasing the concentration of the measured molecules. Different burst analysis parameter values for different concentrations (62, 31 & 15 pM, shown as filled circle, empty circle and asterisk, respectively) of molecules diffusing with a constant diffusion coefficient (90 μm²/s). The relative occurrence of impure bursts was calculated as the fraction of bursts with an impurity level larger than 0, as the fraction of non-single-molecule bursts. The main level of impurity was calculated as the mean of all burst impurity levels, as the mean fraction of photons from other molecules (other than the main molecule in a burst).

Additionally, figure 8 shows that decreasing the concentration of the measured molecules further decreases the molecular position dispersion.

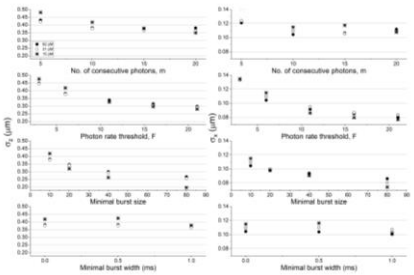


Figure 8. The molecular position dispersion decreases when decreasing the concentration. Different burst analysis parameter values for different concentrations (62, 31 & 15 pM, shown as filled circle, empty circle and asterisk, respectively) of molecules diffusing with a constant diffusion coefficient (90 μm²/s). Quantification of the position dispersion in x, y & z using different burst analysis parameter values. The quantification was achieved via calculation of the standard deviation of the molecule positions in x (right) & z (left). From top to bottom, we quantified the molecules' positional narrowing as a function of a sliding window of *m* consecutive photons, using a constant photon rate threshold of *F*=6; as a function of the photon rate threshold *F*; as a function of the minimal burst size, , for burst search results using *m*=10 & *F*=6; and as a function of the minimal burst width, , for burst search results using *m*=10 & *F*=6.

Overall, and as expected, measurements at lower molecular concentrations give rise to better single-molecule bursts with high purity as single molecules and that are detected when crossing close to the center of the effective detection volume. Analyzing such results with elevated values of photon rate threshold, *F*, and moderate burst size threshold values further improve the quality of the single molecule bursts.

2.3.2. Decreasing the diffusion coefficient further decreases the occurrence of impure single-molecule bursts, but not necessarily the molecular position dispersion

The effect of molecules diffusing through the effective detection volume with different diffusion coefficients on the quality of the single-molecule bursts is a bit less trivial. A molecule that diffuses slowly inside a region of the effective detection volume that is associated with a given photon rate, will move to another region associated with a different photon rate, slowly. Therefore, the rate in which the instantaneous photon rate changes will be slower. This means that for a given photon rate threshold, *F*, once the molecule emits photons with a higher rate, it takes a longer time for it to diffuse into regions in which it will emit photons at a rate lower than *F*. How should the decrease in diffusion coefficient value influence the amount of impure bursts that are selected by the burst analysis, the level of impurity in these bursts, as well as the molecular position dispersion around the center of the PSF? To answer these questions, we compared the molecular position dispersion and the fraction and level of impure bursts, using different simulations with different diffusion coefficient values (90.0, 22.5 & 5.6 μm²/s), at a constant concentration (62 pM). As in the previous assessment, we have tested different burst search analysis parameter values and their performance in reducing the molecular position dispersion, the fraction of impure bursts and the mean impurity levels.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 May 2019 doi:10.20944/preprints201905.0327.v1

Figure 9 shows that when it comes to the amount of impure bursts, as well as the level of impurity, the slower the measured molecules diffuse, the lower will the occurrence of impure bursts be. Additionally, for the leftover impure bursts that passed the burst selection, their level of impurity will be smaller than in faster diffusing molecules.

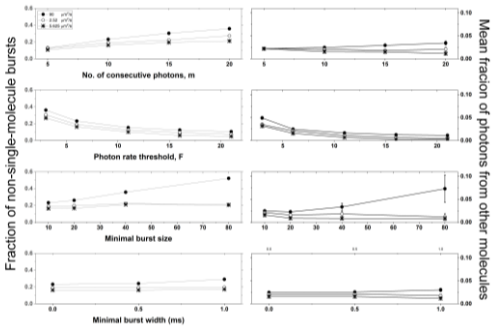


Figure 9. The occurrence and level of impure single-molecule bursts decrease when decreasing the value of their diffusion coefficient. Different burst analysis parameter values of molecules diffusing with a different diffusion coefficient values (90, 22.5 & 5.6 $\mu m^2/s$, shown as filled circle, empty circle and asterisk, respectively) for a constant concentration (62 pM). The relative occurrence of impure bursts was calculated as the fraction of bursts with an impurity level larger than 0, as the fraction of non-single-molecule bursts. The main level of impurity was calculated as the mean of all burst impurity levels, as the mean fraction of photons from other molecules (other than the main molecule in a burst).

Figure 10, however, shows that when it comes to the effect of the diffusion coefficient values on the molecular position dispersion around the center of the PSF, the results are not that straightforward.

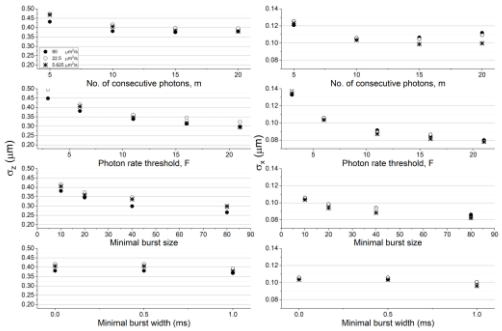


Figure 10. The molecular position dispersion can decrease or increase when decreasing the value of the value of the diffusion coefficient. Different burst analysis parameter values of molecules diffusing with a different diffusion coefficient values (90, 22.5 & 5.6 $\mu m^2/s$, shown as filled circle, empty circle and asterisk, respectively) for a constant concentration (62 pM). Quantification of the position dispersion in x & z using different burst analysis parameter values. The quantification was achieved via calculation of the standard deviation of the molecule positions in x (right) & z (left). From top to bottom, we quantified the molecules' positional narrowing as a function of a sliding window of m consecutive photons, using a constant photon rate threshold of F=6; as a function of the photon rate threshold F; as a function of the minimal burst size, , for burst search results using m=10 & F=6; and as a function of the minimal burst width, , for burst search results using m=10 & F=6.

It seems that reduced value of the diffusion coefficient actually induces a non-monotonic change in the molecular position dispersion. A glimpse at the shape of the distribution of positions in these simulations may help better understand the reason behind the non-monotonic changes in the molecular position dispersion as a function of the value of the diffusion coefficient. Inspecting the shape of the molecule position distribution in figure 2, one can clearly observe an asymmetric-shaped single population distribution. This was the result when stringent burst analysis parameter values were used (m=10, F=6 and a burst size threshold of 40 photons). When using minimal burst search

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 May 2019 doi:10.20944/preprints201905.0327.v1

parameter values (m=5 & F=6; Figure 1), the shape may imply two sub-populations of molecule positions. This might be expected for a physical simulation of the PSF (Figure S1) that is different from the ideal symmetrical Gaussian approximation. Nevertheless, using stringent burst analysis parameters, the position distribution looks as if the two sub-populations have been averaged into one population in a skewed distribution (Figure 2).

We therefore inspected the molecule position distribution for the simulations with slower diffusion coefficients (22.5 & 5.6 μm²/s in Figures S2 & 11, respectively).

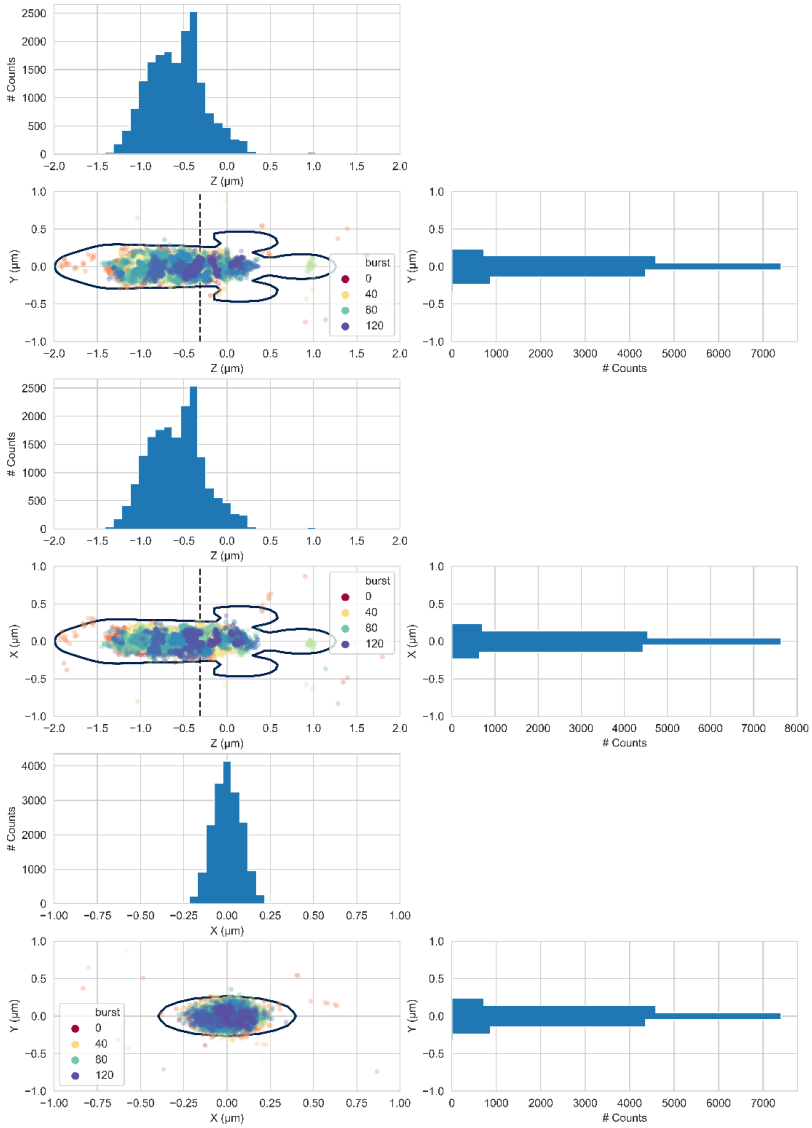


Figure 11. The positions of diffusing molecules when they emitted photons that were detected and selected by the burst analysis, with burst analysis parameters m=10, F=6 & a minimal burst size threshold of 40. In the top, central & bottom panels we show the 2D projections at the yz, xz & xy planes when x=0, y=0 & z=0, respectively. Each dot in the scatter plots is an emitted photon. These results are for the simulation of 15 molecules in 425 fL rectangular box (yielding a concentration of 62 pM), where the diffusion coefficient of the molecules was 5.625 μm²/s. The colors of the points correspond to the burst number out of the overall number of bursts. In each panel, the 1D projections are shown in histograms.

Inspecting molecule position distribution for these simulations, after burst selection using the stringent burst analysis parameter values, we observe that the slower the molecules are diffusing, the more pronounced are the two sub-populations of molecule positions. This feature may reflect motional narrowing. In single-molecule fluorescence detection, the slower molecules diffuse through a non-uniform asymmetric effective detection volume. The slower molecules diffuse, the more

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 May 2019 doi:10.20944/preprints201905.0327.v1

photons they will emit per each region of space. Therefore, the slower molecules diffuse, the better their molecular positions will reveal the fine shape of the effective detection volume. On the other hand, fast diffusing molecules emit a few photons per each region of space, which may lead to averaging out of the molecular positions, and to what is perceived as motional narrowing, observed very well in figure 2. In our case, if indeed there are two distinct local maxima of photon rates in the effective detection volume, the positions at which molecules emitted photons will include both position sub-populations.

The appearance of two sub-populations of molecular positions is what increases the calculated position dispersion, since its calculation is based on the standard deviation of all positions. Obviously, the bimodal position distribution introduces a larger standard deviation. Additionally, the standard deviation of all molecule positions becomes irrelevant as a measure of dispersion, if the distribution of positions is of more than a single population.

Overall, the recommendation that this work provides for experimentalists using confocal-based SMFD is to try keeping the concentrations of the measured molecules as low as feasibly possible, and to analyze the experimental results with a large photon rate threshold and a modest burst size threshold. However, since it is mostly the value of F that controls the molecular position dispersion, caution should be taken in choosing a very large value of F, when measuring molecules with extremely low diffusion coefficients. That is since a large value of F, may introduce two types of bursts, from molecules that traversed two position sub-populations that became distinguishable by the large F value.

2.3. Improving the accuracy of mean FRET efficiency estimation

We have shown that proper choice of burst analysis parameter values can greatly influence the occurrence of impure single-molecule bursts as well as their level of impurity. Next, we show how impure photon bursts may influence measurements based on ratios of burst photon counts.

In applications of confocal-based single-molecule fluorescence detection, histograms of ratios of photon counts are many times the main plots from which inferences are made. For instance, in single-molecule Förster resonance energy transfer (smFRET), FRET histograms are constructed out of FRET efficiencies of photon bursts, after the fluorescence signal is split onto two SPADs, using a dichroic mirror. The FRET efficiency of a burst is calculated by taking the ratio of all photon detected in the acceptor fluorescence detection channel in the numerator and all photons detected both in the donor and acceptor fluorescence detection channels in the denominator. The shape of the histogram of the FRET efficiencies of all bursts can help make inferences on whether the sample included a single population, or distinct sub-populations, based on FRET. This, in turn, is the power of smFRET – it helps define the amount of molecular or conformational species, as well as their mean FRET efficiencies.

Imagine that a sample contains a mixture of two types of molecules, yielding different FRET efficiencies. Assume the mean FRET efficiency of sub-population one is <E>=0.75, and that the mean FRET efficiency of the second sub-population os <E>=0.50. If a molecule from the second sub-population (the one with <E>=0.50) traverses the effective detection volume, it produces both donor and acceptor photons, with a FRET efficiency close to 0.50, with the deviation induced mostly due to lack of enough photons. This deviation is not systematic. Now, imagine that while that molecule traverses the effective detection volume, another molecule belonging to the first sub-population (the one with <E>=0.75), also traverses a part of the effective excitation volume. There is a possibility that it will also emit a few photons, however since <E>=0.75, there is a higher probability that these photons will be detected in the acceptor detection channel. Therefore, <E>=0.75 molecular impurity in a <E>=0.50 burst, systematically biases the FRET efficiencies to values higher than 0.50.

In the previous sections, we have shown how different burst analysis parameter values influence the occurrence and level of burst impurity. More specifically, we have shown how increasing the value of the photon rate threshold, F, reduces both the occurrence of impure bursts and also reduces the level of impurity in the leftover impure bursts. Therefore, we can anticipate that fitting a FRET

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 May 2019 doi:10.20944/preprints201905.0327.v1

histogram of two sub-populations will yield mean FRET efficiencies with values that deviate from the ground-truth simulated values, and that the higher F will be, the smaller these differences will be. We therefore simulated the free 3D diffusion of 15 molecules, at a concentration of 62 pM and with a diffusion coefficient of 90 $\mu m^2/s$. Then we allocated donor and acceptor photon timestamps for 10 out of 15 molecules, according to <E>=0.75, and for the leftover 5 molecules, according to <E>=0.5. Then, we analyzed the results for bursts using a sliding window of m=10 consecutive photons, with different values of the photon rate threshold, F={6, 11, 21}. We collected the FRET efficiencies of all bursts into FRET histograms and fitted them with a sum of two-gaussians' function. The results of this procedure are shown in figure 12.

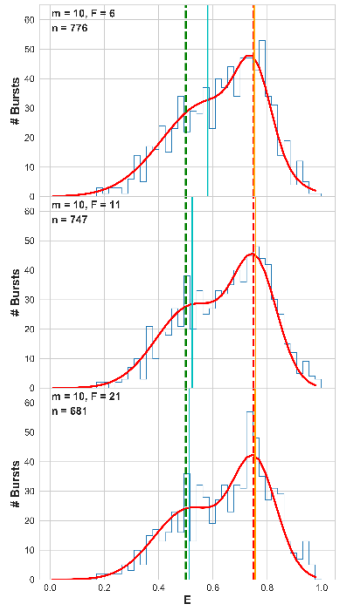


Figure 12. Increasing the value of the photon rate threshold, F, improves the accuracy of the retrieved mean FRET efficiency. From top to bottom, each panel shows the resulting FRET histogram (blue), the best fit sum of two-gaussians (red), the best-fit mean FRET efficiencies (orange and cyan), and the simulation ground-truth mean FRET efficiency values (dashed red and green). These results are for the simulation of 15 molecules in 425 fL rectangular box (yielding a concentration of 62 pM), where the diffusion coefficient of the molecules was 90 $\mu m^2/s$, and the molecules were split to 10 with <E>=0.75 & 5 with <E>=0.5.

Observe how the higher the value of F was, the closer the best fit mean FRET efficiencies were to the ground-truth values. For the sake of explanation, let us focus on the <E>=0.50 sub-population. Once in a while, a molecule that belongs to the <E>=0.75 population, diffuses through parts of the effective detection volume, at the same time a molecule of the <E>=0.50 sub-population traverses the its center. Both molecules emit both donor and acceptor photons. The <E>=0.50 molecule emits donor and acceptor photons with equal probabilities and emit most of the photons. However, also the <E>=0.75 molecule emits a few photons, with a higher probability of them being red. Overall, this burst may systematically include more acceptor photons than donor ones, and hence its FRET efficiency may be higher than 0.50. With enough impure bursts (the higher the occurrence of impure bursts) and high enough level of impurity (the amount of photons not originating from the main molecule), this bias will be large enough to influence the shape of the whole sub-population. This scenario is what can be seen in the top panel of figure 12. However, the panels below show how the higher the value of F is, the smaller this bias becomes, exactly as the assessments made in the text above and in figure 6, predicted.

Additionally, not only the accuracy of the retrieved mean FRET efficiency improves, but also the fraction of the sub-populations. The ground-truth value of the fraction of the <E>=0.75 sub-population was 0.66 (10 molecules out of a total of 15). The values retrieved by the fitting procedure for F=6, 11 & 21 are 0.47±0.07, 0.58±0.04 & 0.61±0.05, respectively. Clearly, using a large F value is a good practice

 doi:10.20944/preprints201905.0327.v1

when the accuracy of the retrieved values of the mean FRET efficiencies and the fraction of the FRET sub-populations is important.

3. Discussion

In this work, we have systematically tested the effect of different burst analysis parameters on the underlying diffusing molecules. Specifically we focused on two main results: 1) how close were the molecules, that were detected as photon bursts, to the center of the effective detection volume, we referred to as the molecular position dispersion, and 2) how often and how well were photon bursts originating from single-molecule. We have shown that increasing the values of burst analysis parameters helps in reducing the molecular position dispersion. However, the parameters that when increasing their values mostly influence the reduction of the molecular position dispersion were the photon rate threshold, *F*, used in the burst search procedure and the burst size threshold used afterwards in burst selection. Then we have shown that increasing the value of *F* helps in reducing the amount of impure bursts and their level of impurity. Increasing the value of the burst size threshold, on the other hand, introduces an increase in burst impurity.

Then, we have shown that measuring the molecules at lower concentrations, help reduce both the molecular position dispersion and burst impurity. However, on the other hand, it might sometimes not be practical. The lower the concentration is, the longer data acquisition will take until a high enough number of selected bursts is achieved. When performing smFRET experiments, if the labeled molecules are at a concentration of 50-100 pM, proper data acquisition of enough legible bursts (bursts that passed the burst search & selection procedure) can take 5-10 minutes. However, decreasing the concentration by an order of magnitude will increase acquisition time by an order of magnitude, which for some applications, and for some experimentalists, may be considered too long for a single measurement. State-of-the-art multisport single-molecule spectroscopy allows to mitigate this problem, by parallelizing a large number of independent SMFD measurements[11].

We have also assessed what would be the effect of decreasing diffusion coefficient values on the resulting single-molecule bursts. Measuring slow diffusing molecules can further help decrease the occurrence and level of impure bursts. However, when it comes to molecular position dispersion, one has to take into account the fact that motional narrowing may make the molecule position distribution appear as a single population for rapidly diffusing molecules. Motional narrowing does not occur for slow diffusing molecules, which makes the molecule position distribution appear somewhat bi-phasic. Therefore, at the extreme of measuring slow diffusing molecules, caution should be taken when choosing high *F* values in the burst analysis, to prevent the separation into two groups of bursts from two local maxima of the effective detection volume. It should be noted, however, that we did not cover all possible diffusion coefficient values. Taking into account that the diffusion coefficients of small proteins such as protein L, BLIP & Adenylate Kinase are ~135, ~80 & ~70 $\mu m^2/s$, respectively[19], and of large complexes such the 70S ribosomal subunit is ~10 $\mu m^2/s$ [20], we believe we covered the value range typical to most biomolecules. Therefore, we believe that the possible effect of no motional narrowing due to very low diffusion coefficients will not be observed in most SMFD measurements of biomolecules.

Finally, we have shown the effect of impure photon bursts on how accurate one can retrieve the mean FRET efficiency from histograms of burst-wise FRET efficiencies. In situations when a mixture of molecular species are measured, bursts of one molecular species that are contaminated by photons arising from molecules of the other species can influence the overall mean FRET efficiency retrieved from fits to the FRET histogram. This can be explained by understanding that molecular species distinguished by different mean FRET efficiency, <E>, values have a different amount of donor and acceptor photons emitted. A high <E> molecule contaminating a burst arising from a low <E>, may overall increase the acceptor photons in its bursts, yielding higher experimental E values. The recommendation to use elevated values of the photon rate threshold, *F*, proved beneficial in restoring the accuracy of the retrieved mean FRET efficiency (Figure 12). This is highly important in the application of smFRET to retrieve information on accurate donor-acceptor distances (reviewed by Lerner & Cordes *et al.*[21]). In this context, it is worth noting that impure photon bursts may also bias

 doi:10.20944/preprints201905.0327.v1

other estimates based on photon counts in a burst, and their ratios, such as, for instance, in single-molecule fluorescence anisotropy. In fluorescence anisotropy, much like in smFRET, the fluorescence signal is split onto two SPADs, only by a polarizing beam splitter, rather than by a dichroic mirror. The fluorescence anisotropy is a ratio of photon counts, and hence is influenced by single-molecule impurity from other molecules with a different mean fluorescence anisotropy. Therefore, the recommendation to use elevated values of *F* in burst analysis of SMFD measurements is of general use when analyzing single-molecule fluorescence bursts of freely diffusing molecules in SMFD measurements.

It is noteworthy to mention that in our survey we did not cover all characterizations of single molecule bursts. For instance, while we focused on molecules freely diffusing in 3D, the next logical steps would be to test the performance of the burst search analysis on 2D diffusion (for membrane proteins), 1D diffusion (for filament-associating proteins), and in situations where not only diffusion occurs but also convection and perhaps flow. These features and others will be the subject of further investigation in future work.

4. Materials and Methods

We performed all 3D diffusion simulations using PyBroMo (python Brownian motion) simulations (https://github.com/tritemio/PyBroMo; was utilized in previous works[13–16]). Afterwards, we simulated the photon timestamps to simulate SMFD of freely diffusing molecules, also using the PyBroMo code. All of the PyBroMo and FRETbursts[18] code used here were documented in Jupyter notebooks, that were deposited in Zenodo[22]. These simulations produced photon HDF5 files, that hold all the simulated molecule positions, photon timestamps and photon identity (for smFRET simulations – either donor or acceptor detection channels), in files with names beginning with 'pybromo_', 'times_' and 'smFRET_', respectively. The photon HDF5 files of the photon timestamps and identities were deposited in Zenodo[23]). The molecule diffusion trajectory photon HDF5 files were not deposited, due their large size. These files can be reproduced by using the '*PyBroMo - 1-. Simulate 3D trajectories - single core - different 3D diffusion simulation conditions.ipynb*' Jupyter Notebook (deposited in Zenodo[22]), where we documented the input parameter values for all simulation conditions performed in this work.

In a nutshell, all photon HDF5 files carrying photon timestamps were analyzed by FRETbursts[18], by first estimating the background rate, and then, by using the sliding window algorithm[6,11,12]. Then, photons from single-molecules were identifying as having instantaneous photon rates larger than *F* times the background rate. Identified bursts, were further selected by using different types of burst selection criteria. Then, per each set of burst analysis results, we tested which was the molecule that produced each burst, its level of impurity (how many photons in the burst originated from molecules other than the main one), and the positions of the molecule when it emitted the burst's photons. All the details are specified in the text above as well as in the Jupyter notebooks deposited in Zenodo[22].

All figures in this work were produced either by *matplotlib* in the Jupyter notebooks, or by OriginLab *Origin 2018*.

Supplementary Materials: Supplementary File 1: manuscript_SI (PDF, 521 KB)

Author Contributions: Conceptualization, E.L.; methodology, E.L.; validation, E.L.; formal analysis, D.H. & E.L.; investigation, D.H. & E.L.; resources, E.L.; data curation, D.H. & E.L.; writing—original draft preparation, E.L.; writing—review and editing, D.H. & E.L.; visualization, D.H. & E.L.; supervision, E.L.; project administration, E.L.; funding acquisition, E.L.".

Funding:

Acknowledgments: We would like to thank Antonino Ingargiola for the assistance and consultation regarding the use of the PyBroMo simulation code. Additionally, we would like to thank Xavier Michalet and Shimon Weiss for fruitful discussions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lerner, E.; Cordes, T.; Ingargiola, A.; Alhadid, Y.; Chung, S.; Michalet, X.; Weiss, S. Toward Dynamic Structural Biology: Two Decades of Single-Molecule Förster Resonance Energy Transfer. *Science (80-.).* **2018**, *359*, eaan1133, doi:10.1126/science.aan1133.
2. Cohen, A. E.; Moerner, W. E. The Anti-Brownian ELectrophoretic trap (ABEL trap): fabrication and software. In *Proc.SPIE*; 2005; Vol. 5699.
3. Boukobza, E.; Sonnenfeld, A.; Haran, G. Immobilization in Surface-Tethered Lipid Vesicles as a New Tool for Single Biomolecule Spectroscopy. *J. Phys. Chem. B* **2001**, *105*, 12165–12170, doi:10.1021/jp012016x.
4. Okumus, B.; Wilson, T. J.; Lilley, D. M. J.; Ha, T. Vesicle Encapsulation Studies Reveal that Single Molecule Ribozyme Heterogeneities Are Intrinsic. *Biophys. J.* **2004**, *87*, 2798–2806, doi:https://doi.org/10.1529/biophysj.104.045971.
5. Rhoades, E.; Cohen, M.; Schuler, B.; Haran, G. Two-State Folding Observed in Individual Protein Molecules. *J. Am. Chem. Soc.* **2004**, *126*, 14686–14687, doi:10.1021/ja046209k.
6. Fries, J. R.; Brand, L.; Eggeling, C.; Köllner, M.; Seidel, C. A. M. Quantitative Identification of Different Single Molecules by Selective Time-Resolved Confocal Fluorescence Spectroscopy. *J. Phys. Chem. A* **1998**, *102*, 6601–6613, doi:10.1021/jp980965t.
7. Ying, L.; Wallace, M. I.; Balasubramanian, S.; Klenerman, D. Ratiometric Analysis of Single-Molecule Fluorescence Resonance Energy Transfer Using Logical Combinations of Threshold Criteria: A Study of 12-mer DNA. *J. Phys. Chem. B* **2000**, *104*, 5171–5178, doi:10.1021/jp993914k.
8. Eggeling, C.; Berger, S.; Brand, L.; Fries, J. R.; Schaffer, J.; Volkmer, A.; Seidel, C. A. M. Data registration and selective single-molecule analysis using multi-parameter fluorescence detection. *J. Biotechnol.* **2001**, *86*, 163–180, doi:https://doi.org/10.1016/S0168-1656(00)00412-0.
9. Lee, N. K.; Kapanidis, A. N.; Wang, Y.; Michalet, X.; Mukhopadhyay, J.; Ebright, R. H.; Weiss, S. Accurate FRET Measurements within Single Diffusing Biomolecules Using Alternating-Laser Excitation. *Biophys. J.* **2005**, *88*, 2939–2953, doi:10.1529/biophysj.104.054114.
10. Nir, E.; Michalet, X.; Hamadani, K. M.; Laurence, T. A.; Neuhauser, D.; Kovchegov, Y.; Weiss, S. Shot-Noise Limited Single-Molecule FRET Histograms: Comparison between Theory and Experiments. *J. Phys. Chem. B* **2006**, *110*, 22103–22124, doi:10.1021/jp063483n.
11. Ingargiola, A.; Lerner, E.; Chung, S.; Panzeri, F.; Gulinatti, A.; Rech, I.; Ghioni, M.; Weiss, S.; Michalet, X. Multispot Single-Molecule FRET: High-Throughput Analysis of Freely Diffusing Molecules. *PLoS One* **2017**, *12*, e0175766, doi:10.1371/journal.pone.0175766.
12. Eggeling, C.; Fries, J. R.; Brand, L.; Günther, R.; Seidel, C. A. M. Monitoring conformational dynamics of a single molecule by selective fluorescence spectroscopy. *Proc. Natl. Acad. Sci.* **1998**, *95*, 1556 LP – 1561, doi:10.1073/pnas.95.4.1556.
13. Ingargiola, A.; Laurence, T.; Boutelle, R.; Weiss, S.; Michalet, X. Photon-HDF5: An Open File Format for Timestamp-Based Single-Molecule Fluorescence Experiments. *Biophys. J.* **2016**, *110*, 26–33, doi:10.1016/j.bpj.2015.11.013.
14. Lerner, E.; Ingargiola, A.; Weiss, S. Characterizing Highly Dynamic Conformational States: The Transcription Bubble in RNAP-Promoter Open Complex as an Example. *J. Chem. Phys.* **2018**, *148*, 123315, doi:10.1063/1.5004606.
15. Ingargiola, A. Simulation of freely-diffusing smFRET measurements 2016.
16. Ingargiola, A. Simulation of freely-diffusing smFRET data of a static mixture of 2 populations 2016.
17. Nasse, M. J.; Woehl, J. C. Realistic modeling of the illumination point spread function in confocal

scanning optical microscopy. *J. Opt. Soc. Am. A* **2010**, *27*, 295–302, doi:10.1364/JOSAA.27.000295.

18. Ingargiola, A.; Lerner, E.; Chung, S.; Weiss, S.; Michalet, X. FRETBursts: An Open Source Toolkit for Analysis of Freely-Diffusing Single-Molecule FRET. *PLoS One* **2016**, *11*, e0160716.
19. Sherman, E.; Itkin, A.; Kuttner, Y. Y.; Rhoades, E.; Amir, D.; Haas, E.; Haran, G. Using Fluorescence Correlation Spectroscopy to Study Conformational Changes in Denatured Proteins. *Biophys. J.* **2008**, *94*, 4819–4827, doi:10.1529/biophysj.107.120220.
20. Kempf, N.; Remes, C.; Ledesch, R.; Züchner, T.; Höfig, H.; Ritter, I.; Katranidis, A.; Fitter, J. A Novel Method to Evaluate Ribosomal Performance in Cell-Free Protein Synthesis Systems. *Sci. Rep.* **2017**, *7*, 46753, doi:10.1038/srep46753.
21. Lerner, E.; Cordes, T.; Ingargiola, A.; Alhadid, Y.; Chung, S.; Michalet, X.; Weiss, S. Toward dynamic structural biology: Two decades of single-molecule Förster resonance energy transfer. *Science (80-.).* **2018**, *359*.
22. Hagai, D. and; Lerner, E. Systematic assessment of freely-diffusing single-molecule fluorescence detection using Brownian motion simulations - Jupyter notebooks that document the photon timestamp simulations as well as their analysis by FRETbursts 2019.
23. Hagai, D. and; Lerner, E. Systematic assessment of freely-diffusing single- molecule fluorescence detection using Brownian motion simulations - photon timetag simulation files 2019.