Simulated Biological Cells for Receptor Counting in Fluorescence Imaging

Julien Ghaye · Giovanni De Micheli · Sandro Carrara

Abstract Digital image processing and epifluorescence microscopy provide one of the main and basic tools for living biological cell analysis and studying. Developing, testing, and comparing those image processing methods properly is eased by the use of a controlled environment. Taking advantage of an existing database of verified and trustworthy images and metadata helps controlling the validity of the processing results. Manually generating that golden database is a long process involving specialists being able to apprehend and extract useful data out of fluorescent images. Having enough cases in the database to challenge the processing methods and gain trust in them can only be achieved manually through time-consuming, prone to human-error processes. More and more we need to automate this process. This paper presents a framework implementing a novel approach to generate synthetic fluorescent images of fluorescently stained cell populations by simulating the imaging process of fluorescent molecules. Ultimately, the proposed simulator allows us to generate images and golden data to populate the database, thus providing tools for the development, evaluation, and testing of processing algorithms meant to be used in automated systems.

Keywords Biological image processing · Fluorescence imaging · Cell simulation · Synthetic image · Simulator · Deconvolution

J. Ghaye (✉)
Integrated Systems Laboratory, EPFL IC ISIM LSI1
INF336 (INF Building) Station 14, 1015 Lausanne,
Switzerland
e-mail: julien.ghaye@epfl.ch

G. De Micheli
Integrated Systems Laboratory, EPFL IC ISIM LSI1
INF341 (INF Building) Station 14, 1015 Lausanne,
Switzerland

S. Carrara
Integrated Systems Laboratory, EPFL IC ISIM LSI1
INF338 (INF Building) Station 14, 1015 Lausanne,
Switzerland

1 Introduction

Fluorescence microscopy is a non-intrusive imaging system that enables to image in situ cells and target some biological details such as the DNA distribution inside the cell nucleus [1] or even the movement of surface receptors [2]. These applications always require the use of digital image processing techniques in order to precisely locate their target and extract valuable information. For example, once the fluorescent signal of a biological target has been recovered, determining quantitative relationships between the signal and biological metrics both in 2D images [3] and in reconstructed 3D structures [4] is possible. These examples show that signal processing for fluorescence microscopy has become important to analyze and extract data. The applied image processing methods have to be tested and validated before being used with enough trust in an automated system. Aside from the classic heuristic algorithms development cycle, which is based on refining processing methods by working on actual real images, simulation tools provide a new approach for working on image processing algorithms. As showed in [5, 7], verifying and validating the information extracted by those algorithms can be done by comparing them to golden meta-data, which are known to be valid and correct.

In other words, processing methods under evaluation are asked to find solutions that are already known to test their capabilities. This approach allows one to quantify and compare realistically different processing methods. Generating that golden database can be done either manually or by using a simulation framework. The former is not convenient for various reasons. Manually analyzing and extracting golden meta-data from a database of images can be time-consuming. It is prone to human error and requires an experimental setup to gather those images. A tool simulating an environment and the imaging system is able to provide both the golden meta-data and the related images. Assuming the simulator is able to embed features in the produced images targeted by some given algorithms, these can be quantitatively compared and studied. In the case of fluorescence imaging for biological applications, the variability expressed by the samples is huge. It is difficult to have enough control over the biological fluorescent samples in order to generate images for the database that exhibit one or many desired features. On the other hand, a simulator does not suffer from this lack of controllability.

The goal of this paper is to present a tool that can be used to develop and evaluate algorithms for fluorescently stained samples, in particular for estimating variations in the amount of cell receptors. This paper presents an approach to generate synthetic images that is conceptually simple. Two distinct engines form the simulator: the emulation of the sample and the simulation of the optical properties of the detection and imaging system. The former is directly dependent on the experiment that the processing method will work on. It generates a distribution of the fluorophores, in space, present in a cell population. The synthetic images are produced by simulating optical properties such as light diffraction within a microscope or sampling effects of the imager.

Related tools have already been developed by Lehmussola et al. [5, 6] and Svoboda et al. [7]. Unlike this work which breaks down the fluorescence simulation down to the fluorophore level, those are using a sum of Perlin's noise functions [8] as a model for simulating the texture of the fluorescently stained cell nuclei and organelles. Similarly, Lehmussola is generating a cell population and generates bi-dimensional fluorescence images while Svodoba focuses on a single cell but extends the output images in the third dimension by using a 3D *point-spread function* (PSF).

This paper is organized as follows: Section 2 explains the simulation strategy at a high level, keeping implementation and mathematical details aside. Right after, Section 3 describes the mathematics used to generate clouds of fluorophores constituting the simulated cell populations. In Section 4, various optical properties (e.g., sample illumination, PSF, imager simulation) that are implemented in the tool are presented. Section 5 focuses on the computational cost of the implemented tool. Section 6 describes and comments the results obtained by applying a deconvolution algorithm on some images generated by our simulator according to known scenarios. Section 7 shows an example of application of the tool, quantitatively analyzing the capabilities of the deconvolution algorithm with various *signal-to-noise ratio* (SNR) for a given scenario. Finally, the Section 8 concludes this paper.

2 Simulation Strategy

The strategy used for this simulator is intuitive and matches a real fluorescence microscopy setup. On one hand, we have a fluorescently stained biological sample. On the other hand, we have the imaging system (i.e., the optics and image sensor couple).

First, we need to emulate a fluorescent sample. Such samples involve one or more fluorescent dyes labeling structures and/or molecules of interests. In other words, a fluorescent sample can be seen cloud of fluorescent molecules dispersed in space under the imaging system objective. Labeling a sample with a fluorescent dye implies that a single target (e.g., antibody) can be labeled by one or more fluorophores. Thus, for our simulator, the smallest element we are dealing with is a cluster of fluorophores having a given location in space. One simply has to provide the distribution of the fluorescent clusters in space by modeling the samples to study.

Our work partly deals with image processing methods evaluating variations in the amount of cell receptors within a population of cells. For this purpose, we generate a population of cells using parametric equations. Each cell constitutes of a group of randomly distributed fluorescent clusters, each containing one or many fluorophores. The idea here is to simulate what is happening in a biological sample when receptors are expressed and start grouping together [9]. These clusters, with their location and the amount of fluorophores they embed, represent the basic golden information that processing algorithms can recover.

Second, given the cloud of fluorescent clusters, the framework simulates the imaging process. Each fluorescent cluster is assumed to be a single point of light, emitting photons at a given intensity and wavelength. The tool then simulates the passage of these photons in the optical path. Last, an image sensor, which gathers the photons and samples them to form an

image, is simulated. The optical behavior simulated by the tool is that of a classical wide-field epifluorescence microscope with charge coupled device (CCD) or complementary metal oxide semiconductor imaging. Since fluorescent clusters have a range of sizes within the nanometer scale (below the optical resolution of wide-field microscope), each of them is considered as a single source point of light. Using models of the PSF of the optical system [10], a diffraction pattern of the fluorescent clusters can be computed. Later on, the diffraction pattern is sampled by simulating the behavior of an imager. The output of the latter is the synthetic fluorescent image of the cell population.

Fig. 1 Caco-2 human cells with fluorescently stained TLR-2 receptors

3 Cell Population Simulation

The first task of our simulation tool is the generation of a cloud of fluorescent clusters defining the cell population. The methodology applied to generate this cloud is separated in two distinct steps, using a top-down approach, from population properties to clusters location. First we are generating a population of cells with random locations and shapes based on the parametric models described in [5]. Second, each cell within the generated population is assigned to multiple fluorescent clusters.

3.1 Cell Population

In order to simulate a cell population, we must first define a region where the cells will be distributed. For this, a rectangular area is defined using 2D Cartesian coordinates

$$
\begin{cases}
x \in [0, s_x] \\
y \in [0, s_y]
\end{cases}
\tag{1}
$$

Within this range, a chosen number of points are uniformly distributed using

$$
\begin{bmatrix} x_c \\ y_c \end{bmatrix} = \begin{bmatrix} \mathscr{U}(0, s_x) \\ \mathscr{U}(0, s_y) \end{bmatrix}
\tag{2}
$$

where $\mathscr{U}(a, b)$ defines a uniform distribution within on the interval $[a, b]$. Each of those points $[x_c\ y_c]^T$ is defining the location of the center of a cell cluster constituting the cell population. Furthermore, each cluster is associated with a number of cells also chosen uniformly and randomly within a user-defined range using $\mathscr{U}(c_{min}, c_{max})$. This gives the basic skeleton of the simulated cell population. Grouping cells by clusters comes from the observation that cells are often spatially packed. Figure 1 provides an example of this behavior.
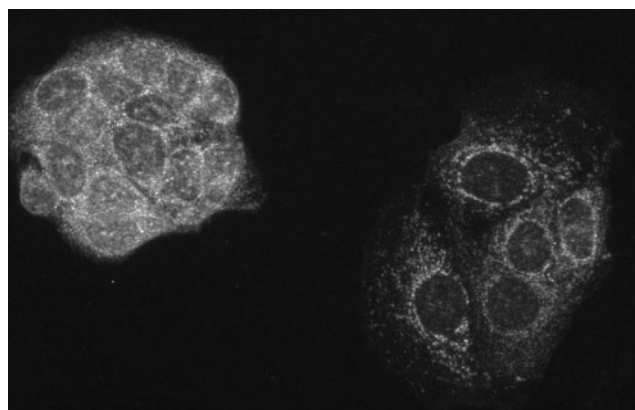
In order to distribute cells around a given cluster center, the simulation engine is using

$$
\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_c \\ y_c \end{bmatrix} + \mathscr{N}(0, \sigma_c^2)
\tag{3}
$$

where $[x\ y]^T$ are the spatial coordinates of a new cell, $[x_c\ y_c]^T$ is the cell cluster center, and $\mathscr{N}(0, \sigma_c^2)$ is a normal distribution with a zero mean and a standard deviation of σ_c.

Once a cell location has been generated, we are using the model presented in [5] to generate the random shape of the nuclei. Using the parametric equation of a circle, we randomize the polar coordinates of the vertices of a regular polygon with k edges. The radial coordinate and the angle coordinate of each vertex ($i = 0, \cdots, k$) are uniformly randomized with

$$
\begin{bmatrix} r_i \\ \theta_i \end{bmatrix} = \begin{bmatrix} r + \mathscr{U}(-\alpha, \alpha) \\ \dfrac{2\pi i}{k} + \mathscr{U}(-\beta, \beta) \end{bmatrix}
\tag{4}
$$

These vertices are then simply connected together using cubic spline interpolation, giving the final shape of the cell nuclei defined as $r_n(\theta)$.

Controlling the overlapping between cells is also an element that must be taken into consideration. While implementing a method that can segment spread cells on a fluorescent image is easy, doing so when cells are packed is another problem. To simulate this, each generated cell is given a radius R with respect to its center whose value is defined by the uniform distribution $\mathscr{U}(R_{min}, R_{max})$. Eventually, if smaller than the maximum value of $r_n(\theta)$, this radius is increased to not cut through the associated nuclei. We define an

Fig. 2 Simulated cells; fluorescent cluster locations on the *left* and associated simulated image on the *right*
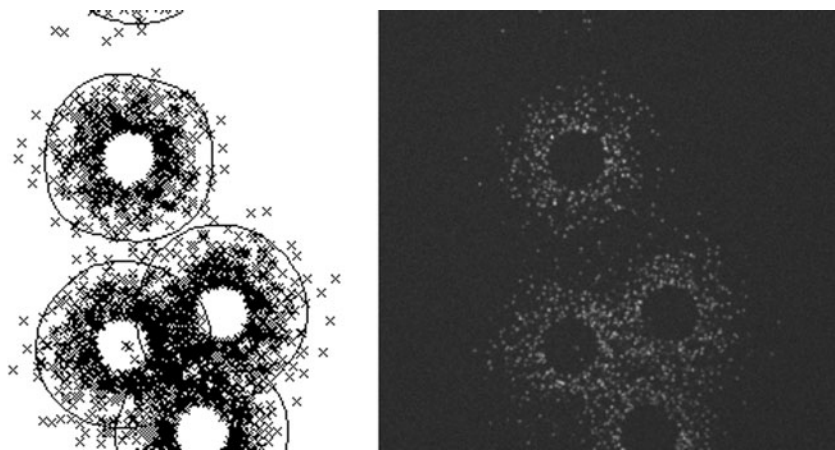
overlapping metric between two synthetic cells *a* and *b* using

$$o_{ab} = \frac{d_{ab}}{R_a + R_b} \tag{5}$$

where d_{ab} is the distance between the two cell centers and $\{R_a, R_b\}$ are their radii. The user can specify the overlapping value above which a new cell can be included in the population.

3.2 Fluorescent Clusters

A fluorescent cluster represents the most simple and basic element in our simulation tool. The fluorescent cluster distribution model we implemented fits images like the one depicted by Fig. 1. We are aiming at modeling cells whose membranes have been permeabilized allowing fluorescently stained clusters of surface receptors to enter the cytoplasm [11]. Modeling the dispersion of these clusters for each cell is done using

$$\begin{bmatrix} r(\theta) \\ \theta \end{bmatrix} = \begin{bmatrix} r_n(\theta) + |\mathcal{N}(0, \sigma_f^2)| \\ \mathcal{U}(0, 2\pi) \end{bmatrix} \tag{6}$$

where $[r\ \theta]^T$ are the polar coordinates of a new cluster location with respect to the cell center and r_n being the parametric equation defining the nucleic shape.

So far, we have only been modeling cell shapes and fluorescent clusters distribution in a bi-dimensional world. In order to take into account from focus fluorophores, we simply add a third dimension toward the simulated optical axis. Each fluorescent cluster is assigned a depth parameter z which is uniformly distributed according to $\mathcal{U}(z_0, z_1)$ where z_0 and z_1 represent the deepest and highest plane, respectively, in between which the clusters will be distributed, with the special case of $z = 0$ representing the in focus plane.

Figure 2 illustrates some synthetic cells that have been generated with the presented simulation engine.

4 Cell Population Imaging

This section accounts for the optical and measurement errors that arise when imaging a fluorescent sample. The fluorescent proteins produce emission light upon excitation; the optical system gathers that light and focuses it onto the imager. Our simulator takes into consideration different errors to mimic the optical system. The first is the uneven exciting light [12], illuminating the cells with a spatially varying intensity. The second is the light diffraction occurring in any optical system [13], which can be fully described by the PSF of the microscope. At last, we model the effects of the imager, which is sampling the image coming out of the simulated microscope and adding various noises such as the photon-shot noise, thermal noise, and readout noise [14]. Figure 3 shows side by side real and simulated clusters of fluorophores where the simulator parameters have been tweaked to match the experimental setup.

4.1 Excitation Light Model

In a classical fluorescence microscopy setup, the source of light cannot be evenly spread onto the sample, creating a vignetting effect of the final images. A quadratic polynomial can be used to model the distribution of light intensity as suggested in [12]. An example of this effect is shown by Fig. 4.

4.2 Optics Modeling

A fluorescent wide-field microscope is an optical system that can be fully modeled by its 3D PSF. The PSF
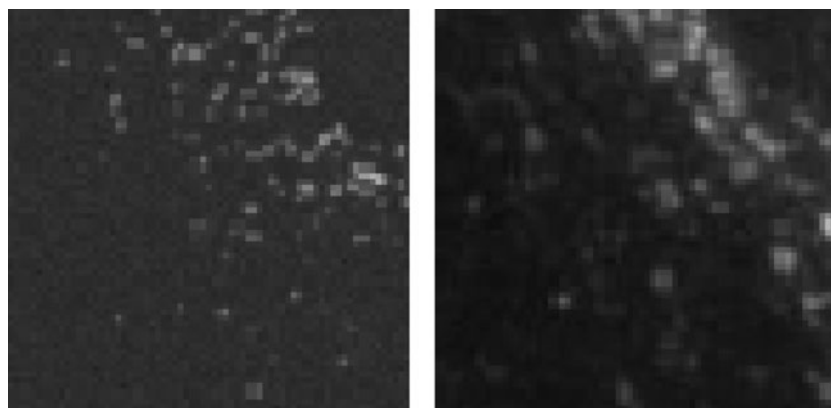
Fig. 3 Comparison of simulated (*left*) and real (*right*) fluorescent clusters. Caco-2 human cells with permeabilized membrane and stained TLR-2 receptors have been used. The *top right of each image* represents a cell nuclei while the fluorescent clusters, bound to the TLR-2 receptors, are emitting photons within the

cell cytoplasm. The real image has been taken with a ×40 air objective having a numerical aperture of 0.9. A CCD imager having 4.4 × 4.4 μm have been used. The simulator has been configured to match the experimental parameters. Each image covers an area of 10 × 10 μm on the sample

of an optical system is the diffraction pattern that is produced when the system is imaging a point source object located anywhere in the object space. In other words, it is its impulse response. Multiple models exist to represent the 3D PSF, among which the Gibson and Lanni scalar model [13] and some Gaussian approximations derived by Zhang et al. in [15]. Due to the amount of fluorescent clusters we might simulate and thus the amount of PSF evaluations the simulator needs to compute, the former model is too computationally intensive. In this implementation of the simulator, we decided to use the least-square Gaussian approximations of the 3D PSF for wide-field microscopes:

$$\text{PSF}(r, z|z_p) = A \exp\left(-\frac{r^2}{2\sigma_\rho^2} - \frac{z^2}{2\sigma_z^2}\right) \qquad (7)$$

where the Gaussian parameters σ_ρ and σ_z are set such that the least-square error is minimized. Their expres-

sion, assuming a wide-field microscope in the paraxial and nonparaxial cases, can be found in [15]. B. Zhang et al. also evaluated the error made by such approximations. According to their conclusion, when the imaged point source is close or in the focal plane, the PSF approximation is accurate. However, it loses accuracy for out-of-focus point sources. We must be aware of this fact while analyzing simulation results.

4.3 Imager Modeling

This part of the simulator is the one generating the final image. It is split into two steps: sampling and acquiring.

The sampling step takes into account the diffraction pattern of each fluorophore cluster, the quantum efficiency of the photosites, the background noise, and the uneven shading presented in Section 4.1. Note that our simulator considers each pixel to be identical [16], i.e., we assume they all have the same quantum efficiency, the same dark current generation,...The diffraction patterns are computed as presented in Section 4.2 assuming no magnification in the simulated microscope. In order to take into account a radial magnification factor, one must find a linear relation between a photosite on the ith line and the jth column and the spatial coordinate in the image space $[x \ y]^T$

$$\begin{bmatrix} i \\ j \end{bmatrix} \leftrightarrow \begin{bmatrix} \dfrac{x}{M} \\ \dfrac{y}{M} \end{bmatrix} \qquad (8)$$

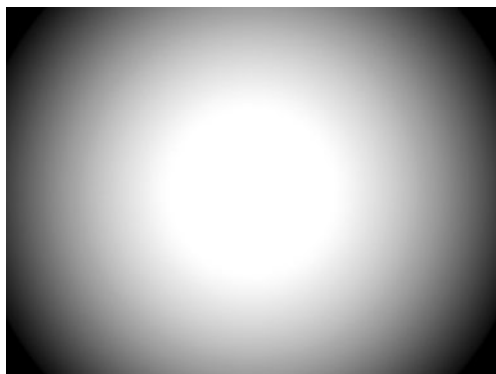where M is the magnification factor specified on the objective. Knowing this relation, we can evaluate the

Fig. 4 Simulated image of an overly exposed sample exhibiting the uneven lighting distribution onto the sample

amount of electrons per second hitting a given photo-site using

$$\frac{\mathrm{d}E}{\mathrm{d}t}(i(x),\, j(y)) = \left[(I * \mathrm{PSF})\, Q_\mathrm{e} + N_\mathrm{g} \right] s(x,\, y) \tag{9}$$

where $I \times \mathrm{PSF}$ is a convolution operation between the ideal image (impulses mapped with the location of the fluorophores) and the PSF, Q_e is quantum efficiency of the photosites, N_g is the constant background noise, and $s(x,\, y)$ models the uneven shading.

The acquiring step simply integrates $\mathrm{d}E(i,\, j)/\mathrm{d}t$ over time in order to determine how many electrons hit a given photosite:

$$N_\mathrm{E}(i,\, j) = \mathscr{P}\left(\int_0^T \left[((S.E(i,\, j)) + N_\mathrm{dc})\, \mathrm{d}t \right] \right) + N_\mathrm{ro} \tag{10}$$

where S can be either 0 or 1 and represents the effect of a shutter, N_dc is the dark current expressed in electron per second, t is the exposure time, N_ro is the readout noise modeled by a normal distribution as suggested in [14], and the operator $\mathscr{P}$ is the photon shot noise. The values of the pixels in the final image are simply derived from N_E using a gain factor expressed in unit per electrons.

5 Computational Cost

The current implementation of the simulator is divided in two parts, generating a cloud of fluorescent clusters and imaging those clusters, as presented in Sections 3 and 4 respectively.

The computational effort of the procedure generating the fluorescent clusters has a linear dependency with the number of clusters to be generated. This was expected according to the process introduced in Section 3, which consists of randomizing the location of the clusters in a defined space.

The imaging process as presented in Section 4 has two major input parameters that influence the execution time. The number of fluorescent clusters to be imaged, and the density, or size of the simulated imager photosites. For each fluorescent cluster, an evaluation of its diffraction pattern contributing to the final image is computed using Eq. 7. This results in a linear time dependency between the execution time and number of clusters. The execution time measurements, graphically depicted in Fig. 5, exhibit this linear behavior.

Furthermore, the execution time of the imaging process has a quadratic dependency with the inverse of the square footage of the simulated imager photosites. Measurements of this behavior are graphically depicted
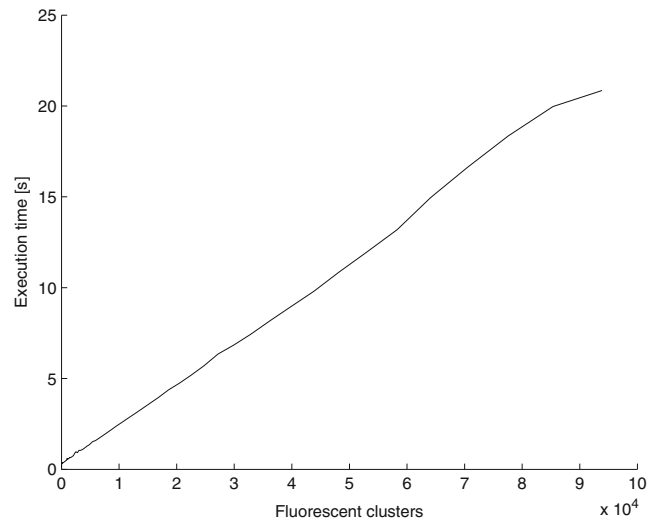

Fig. 5 Execution time of the simulation tool versus the amount of fluorescent clusters in the cell population. Assuming all the other parameters are constant, the time required to image a cloud of fluorescent clusters shows a linear dependency with the amount of clusters. The simulated imager features an array of 1,200 × 1,600 photosites each having a size of 4.4 × 4.4 μm

in Fig. 6. When a fluorescent cluster is imaged by the simulator, Eq. 7 is used for modeling the diffraction pattern and needs to be sampled. Given the location of an imaged cluster on the simulated imager, this equation needs to be evaluated for each photosite surrounding the projected center. This results in a quadratic
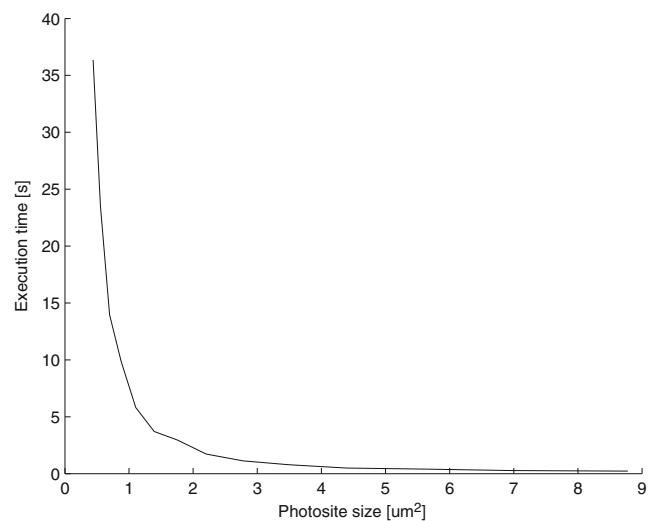

Fig. 6 Execution time of the simulation tool versus the size of the simulated imager photosites. Assuming all the other parameters are constant, the time required to image a cloud of fluorescent clusters shows a quadratic dependency with the inverse of the photosite size. The number of fluorescent clusters was fixed to 1,000. The simulated imager features a constant size active area of 5,280 × 7,040 μm where the photosites are located

behavior. For example, doubling the photosite density by reducing their square footage results in a quadrupled number of evaluations of Eq. 7.

The time required for the addition of noise, as defined by Eqs. 9 and 10, is linearly dependent with the number of photosites in the simulated imager. It is, as expected, independent from the number of fluorescent clusters and is negligible compared to the evaluation of the convolution operation in Eq. 9.

The simulations used to generate Figs. 5 and 6 were executed on a single core of a workstation featuring Intel Xeon X5650 processors at 2.66 GHz using MAT-LAB (R2009b, 64bit).

6 Validation

6.1 Deconvolution Method

In order to validate our image simulator, we are using an existing algorithm that implements means to deconvolve the number of fluorophores present in a fluorescent cluster [17]. Hereunder, we refer to this method as the reference method. It is using basis histograms, like the ones shown in Fig. 7, to process an image gray-level histogram such as the one plotted in Fig. 8. This reference method has been developed and tested by generating intensity histogram directly based on mathematical models according to some parameters (e.g., the number of fluorescent clusters and the amount of fluorophores per cluster). Images have been generated using our simulator based on these parameters. For each scenario, the results produced by the reference method using the mathematically generated histogram and the histogram generated from the simulated images are compared (Table 1).
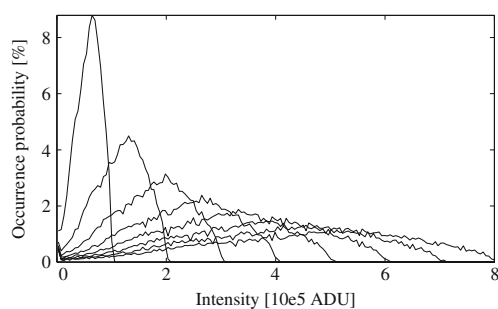

Fig. 8 Example of fluorophore intensity distribution generated from our imager simulator. This distribution represents the simulation scenario 1MD as showed in Table 1

The reference method developed by Mutch et al. [17] is based on their observation showing that in a fluorescent image, the intensity distribution of the pixels containing the useful signal (i.e., the region of interest) is best fitted using a lognormal distribution (probability distribution of a random variable whose logarithm is normally distributed). In that research, the measured intensity of a fluorescent cluster is modeled using

$$I(x, y) = \left(\prod_{j=1}^{J} F_j(x, y)\right) cI \tag{11}$$

where c is the number of fluorophores in the cluster, I is the intensity of a single fluorophore, and F_j represents various effects of the measurement process like the PSF. From the central limit theorem, the intensity distribution of various fluorescent clusters for a large J dictates a lognormal profile.

The various scenarios considered to validate the results computed by the reference method based on

Fig. 7 Basis histograms used to deconvolve intensity distributions. Each curves is a signature of a fluorescent cluster having a single (*high* and *narrow curve*) or multiple fluorophores (*lower* and *wider curves*). It is the quantum information used by the deconvolving method to process histograms
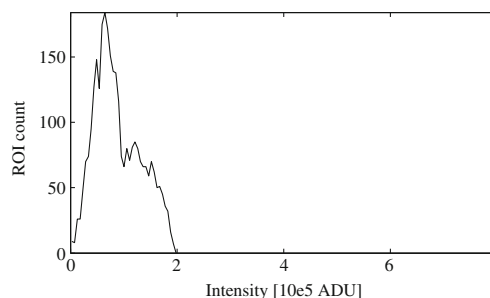
Table 1 Deconvolution results of various simulation cases

Sim. scenario		Sim. value	$\hat{a}_c$ Mutch et al. [17]	$\hat{a}_c$ Our simulator
1MD	a_1	1,500	$1,505 \pm 44$	$1,459 \pm 377$
	a_2	1,500	$1,457 \pm 93$	$1,368 \pm 382$
3MD	a_3	1,500	$1,461 \pm 194$	$1,783 \pm 133$
	a_4	1,500	$1,472 \pm 259$	$1,218 \pm 80$
4MD	a_4	1,500	$1,448 \pm 390$	$2,175 \pm 101$
	a_5	1,500	$1,370 \pm 522$	830 ± 131
6MD	a_4	3,000	$2,703 \pm 320$	$2,568 \pm 987$
7MD	a_1	200	190 ± 46	58 ± 30
	a_2	1,200	$1,205 \pm 134$	$1,345 \pm 71$
	a_3	1,200	$1,226 \pm 288$	$1,350 \pm 147$
	a_4	400	310 ± 267	216 ± 143

The computed estimates are expressed in the form of $\mu \pm \sigma$, where μ is the mean value and σ is the standard deviation over 100 deconvolved intensity profiles

images generated by our simulator are presented in [17] (Table 5) and redisplayed in Table 1. Each scenario has some input parameters indicating the amount of fluorescent clusters a_c having c fluorophores. For example, the scenario labeled 1MD defines a sample having 3,000 fluorescent clusters including 1,500 clusters with 1 fluorophore (a_1) and 1,500 clusters with 2 fluorophores (a_2). From the input parameters a_c, we are generating a couple of intensity profiles for each scenario within Table 1, which are then fed to the reference method for processing. The goal of this method is to find an estimation $\hat{a}_c$ of the a_c.

Each simulation scenario requires two intensity profiles, one generated directly using mathematical models, which is the approach employed in [17] to validate their algorithm, and the other computed from the image generated using the input parameters a_c by our synthetic image generator. The results of the deconvolution of those intensity profiles are presented in Table 1, where the first column of the results indicates the estimate $\hat{a}_c$ computed from the mathematically crafted intensity profiles and the second column shows the estimate $\hat{a}_c$ computed from the intensity profiles of the generated synthetic images. Note that for the latter, a noise removal step is required before deconvolving since the various noises described in Section 4.3 introduce an offset in the intensity profile that must be removed to avoid misresolved fluorophore amounts.

The consistency of the synthetic image generator is established by pointing out that both approaches lead to statistically equal estimates, using the processing algorithm of Mutch et al. [17] as a reference tool and knowing it has already been used in practical applications [18].

6.2 Signal-to-Noise Characteristics

The SNR is a very important parameter to evaluate the quality of an image. Algorithms processing fluorescent images require images either taken with a short exposure time (e.g., Lucky imaging [19]) or with a long exposure time (e.g., super-resolution methods [20]). In those cases, the SNR is either read-noise-limited or photon-noise-limited. A short exposure time favors the influence of the readout noise since the amount of electrons gathered by the photosites from the signal is low. For longer exposure times, the amount of electrons gathered by the photosites is bigger and the readout noise is neglectable compared to the photon-shot noise.

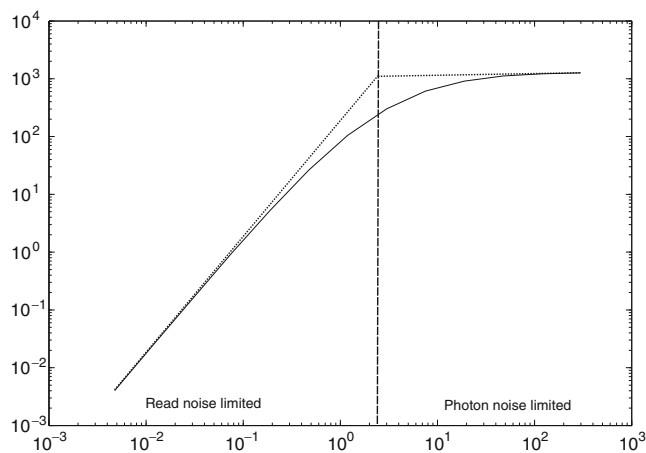Figure 9 shows that images produced by our simulator have this behavior. The plot of the SNR versus the

Fig. 9 SNR versus exposure time. We can distinguish two typical regions: the read-noise-limited region and the photon-shot-noise-limited region

exposure time exhibits the two aforementioned noise-limited regions. The used SNR definition is

$$\text{SNR} = \frac{\sum_{j=1}^{N} \sum_{i=1}^{M} \hat{N}_E(i, j)^2}{\sum_{j=1}^{N} \sum_{i=1}^{M} \left(\hat{N}_E(i, j) - N_E(i, j) \right)^2} \quad (12)$$

$$\text{SNR}_{dB} = 10 log_{10}(\text{SNR}) \quad (13)$$

where $\hat{N}_E(i, j)$ is the expected amount of electrons gathered by the photosite from the signal on the ith line and jth column and N_E is the same, but including electrons generated by noise sources.

7 Application

Evaluating, comparing, and testing image processing algorithms for fluorescence microscopy is directly de-

Table 2 Deconvolution results for the scenario 1MD with various SNR levels

SNR (dB)	$\hat{a}_1$	$\hat{a}_2$
1	0 ± 0	0 ± 0
2	0 ± 1	0 ± 4
3	744 ± 148	$1,104 \pm 209$
4	821 ± 62	$1,299 \pm 64$
5	940 ± 56	$1,661 \pm 99$
6	904 ± 62	$1,611 \pm 130$
7	$1,250 \pm 52$	$1,029 \pm 94$
8	$1,322 \pm 63$	$1,341 \pm 92$
9	$1,313 \pm 57$	$1,522 \pm 112$
10	$1,319 \pm 48$	$1,613 \pm 94$

Expected value for a_1 and a_2 is 1,500. The computed estimates for a given SNR are expressed in the form of $\mu \pm \sigma$, where μ is the mean value and σ is the standard deviation over 100 deconvolved intensity profiles, minus the outliers

pendent on the experiment being performed. Thus, in order to use simulated images for those purposes, the tool must be configured to match the experimental setup such as the optics, imager characteristics, or light illumination. Equally important, the generated cloud of clusters and the number of fluorophores they embed should match the experiment as well.

As an example of application of the developed simulator, we are quantitatively analyzing the capabilities of the deconvolution algorithm [17] for the scenario 1MD by varying the SNR of the produced images. For this analysis, 100 photon-noise-limited images having a fixed average SNR were processed by the deconvolution algorithm. After filtering out the outliers in the results, a couple estimates are computed, for the number of fluorescent clusters having 1 fluorophore ($\hat{a}_1$) and 2 fluorophores ($\hat{a}_2$). The computed estimates for various levels of SNR are gathered in Table 2.

The true, expected values are $a_1 = 1{,}500$ and $a_2 = 1{,}500$ as defined by the scenario 1MD. In those conditions, considering the photon-shot noise and a SNR defined by Eq. 13, we can point that a SNR below 7 dB induce an error on the estimates of over 20% compared to the expected values.

8 Conclusion

This paper presented a new approach to generate synthetic fluorescent images based on imaging fluorescent clusters. The engine is composed of two main parts. First, the location of the fluorescent clusters as well as the number of fluorophores per cluster are generated. Second, these clusters are rendered using models of the imaging system PSF and simulating the sampling properties of an imager. The user is asked for experimental parameters that can be easily found, such as the sample and cell sizes, optical parameters (numerical aperture, magnification,...), and imager specifications (photosite amount and size, quantum efficiency,...).

This tool is primary used as a development tool for image processing algorithms. For example, algorithms whose jobs are to quantitatively evaluate the amount of cell receptors. It falls within the range of simulation tools that are often used for development and debugging purposes in many disciplines. Having a generic and configurable model of a microscope and image sensor for fluorescence microscopy helps with the development of image processing algorithms. By defining the geometry and the location of fluorescent clusters, one can simulate various scenarios and evaluate the effects of the optical properties of the system. This gives us means to simulate a given experiment without having

to set it up in a lab and gather some early intel on how the processing methods are performing in these cases. The simulation tool provides a totally controlled environment where repeatability and border cases can be easily created, which is not always the case with real experiments, especially when dealing with biological samples.

On the other hand, this paper uses the synthetic image generator with a simple distribution model of fluorescently stained cell receptors. The simplicity of this model was a deliberate choice since the goal of this work was to validate the image generator using a know algorithm processing images of fluorescent clusters. Future work would be to extend and try various cell receptor distribution models in order to develop algorithms that can extract quantitative information data about biological samples.

References

1. Mascetti, G., Vergani, L., Diaspro, A., Carrara, S., Radicchi, G., Nicolini, C. (1996). Effect of fixatives on calf thymocytes chromatin as analyzed by 3D high-resolution fluorescence microscopy. *Cytometry, 23*(2), 110–119.
2. Anderson, C. M., Georgiou, G. N., Morrison, I. E., Stevenson, G. V., Cherry, R. J. (1992). Tracking of cell surface receptors by fluorescence digital imaging microscopy using a charge-coupled device camera. Low-density lipoprotein and influenza virus receptor mobility at 4 degrees C. *Journal of Cell Science, 101*(2), 415–425.
3. Mascetti, G., Carrara, S., Vergani, L. (2001). Relationship between chromatin compactness and dye uptake for in situ chromatin stained with DAPI. *Cytometry, 44*(2), 113–119.
4. Nicolini, C., Carrara, S., Mascetti, G. (1997). High order DNA structure as inferred by optical fluorimetry and scanning calorimetry. *Molecular Biology Reports, 24*(4), 235–246.
5. Lehmussola, A., Ruusuvuori, P., Selinummi, J., Huttunen, H., Yli-Harja, O. (2007). Computational framework for simulating fluorescence microscope images with cell populations. *IEEE Transactions on Medical Imaging, 26*(7), 1010–1016.
6. Lehmussola, A., Selinummi, J., Ruusuvuori, P., Niemisto, A., Yli-Harja, O. (2005). Simulating fluorescent microscope images of cell populations. In *27th annual international conference of the engineering in medicine and biology society, 2005. IEEE-EMBS 2005* (pp. 3153–3156).
7. Svoboda, D., Kašík, M., Maška, M., Hubenỳ, J., Stejskal, S., Zimmermann, M. (2007). On simulating 3D fluorescent microscope images. In *Computer analysis of images and patterns* (pp. 309–316).

8. Perlin, K. (1985). An image synthesizer. *ACM SIGGRAPH Computer Graphics, 19*(3), 287–296.
9. Yaqoob, P. (2009). The nutritional significance of lipid rafts. *Annual Review of Nutrition, 29*, 257–282.
10. Haeberlé, O. (2003). Focusing of light through a stratified medium: A practical approach for computing microscope point spread functions. Part I: Conventional microscopy. *Optics Communications, 216*(1–3), 55–63.
11. Jamur, M. C., & Oliver, C. (2010). Permeabilization of cell membranes. *Methods in Molecular Biology, 588*, 63–66.
12. Klein, A., van den Doel, R., Young, I. T., Ellenberger, S., van Vliet, L. (1998). Quantitative evaluation and comparison of light microscopes. In *Proc. SPIE, progress in biomedical optics, optical investigation of cells in vitro and in vivo* (Vol. 3260, pp. 162–173).
13. Frisken-Gibson, S., & Lanni, F. (1991). Experimental test of an analytical model of aberration in an oil-immersion objective lens used in three-dimensional light microscopy. *Journal of the Optical Society of America A, 8*(10), 1601–1613.
14. Mullikin, J. C., van Vliet, L. J., Netten, H., Boddeke, F. R., Van der Feltz, G., Young, I. T. (1994). Methods for CCD camera characterization. In *Proceedings of the SPIE image acquisition and scientific imaging systems* (Vol. 2173, pp 73–84).
15. Zhang, B., Zerubia, J., Olivo-Marin, J. C. (2007). Gaussian approximations of fluorescence microscope point-spread function models. *Applied Optics, 46*(10), 1819–1829.
16. Bigas, M., Cabruja, E., Forest, J., Salvi, J. (2006). Review of CMOS image sensors. *Microelectronics Journal, 37*(5), 433–451.
17. Mutch, S. A., Fujimoto, B. S., Kuyper, C. L., Kuo, J. S., Bajjalieh, S. M., Chiu, D. T. (2007) Deconvolving single-molecule intensity distributions for quantitative microscopy measurements. *Biophysical Journal, 92*(8) 2926–2943.
18. Mutch, S. A., Kensel-Hammes, P., Gadd, J. C., Fujimoto, B. S., Allen, R. W., Schiro, P. G., et al. (2011). Protein quantification at the single vesicle level reveals that a subset of synaptic vesicle proteins are trafficked with high precision. *The Journal of Neuroscience, 31*(41), 1461–1470.
19. Cronin, B., de Wet, B., Wallace, N. I. (2009). Lucky imaging: Improved localization accuracy for single molecule imaging. *Biophysical Journal, 96*, 2912–2917.
20. Huang, B., Bates, M., Zhuang, Z. (2009). Super-resolution fluorescence microscopy. *Annual Review of Biochemistry, 78*, 993–1016.