A Fast Automatic Colocalization Method for 3D Live Cell and Super-Resolution Microscopy

Frédéric Lavancier, Thierry Pécot, Liu Zengzhen, Charles Kervrann

A Fast Automatic Colocalization Method for 3D Live Cell and Super-Resolution Microscopy [1]

*Frédéric Lavancier[1,2], *Thierry Pécot[1], Liu Zengzhen[3], Charles Kervrann[1]

[1] Inria, Centre Rennes-Bretagne Atlantique, SERPICO Project Team, 35042 Rennes, France

[2] University of Nantes, Laboratoire de Mathématiques Jean Leray, 2 rue de la Houssinière, 44322 Nantes, France

[3] Institut Curie, PSL Research University, CNRS UMR 144, Space Time Imaging of Endomembranes Dynamics Team, 75248, Paris, France

Abstract

Colocalizing two fluorescent-labeled proteins remains an open issue in diffraction-limited microscopy and raises new challenges with the emergence of super-resolution imaging, single molecule tagging (PALM, dSTORM...) and high content screening. Two distinct colocalization approaches are usually considered to address this problem : the intensity-based methods are very popular but are known to be sensitive to high intensity backgrounds and provide errors if the signal-to-noise ratio (SNR) is low ; the object-based methods analyze the spatial distribution of the two sets of detected spots by using point process statistics but unfortunately get rid of valuable information by reducing objects to points. We propose a unique method (GcoPS : Geo-coPositioning System) that reconciles intensity-based and object-based methods for various applications in both conventional diffraction-limited and super-resolution microscopy. Unlike previous methods, GcoPS is very fast, robust-to-noise and versatile since it efficiently handles 2D and 3D images, variable signal-to-noise ratios (SNR) and any kind of cell shapes and sizes. The experimental results demonstrate that GcoPS unequivocally outperforms the best competitive methods in adverse situations (noise, chromatic aberrations, ...). The method is able to automatically evaluate the colocalization between large regions and small dots and to detect significant negative colocalization. Since the one-parameter (p-value) GcoPS procedure is very fast in 2D and 3D, it should greatly facilitate objective analysis in large-scale high-content screening experiments.

Keywords : colocalization, hypothesis test, spatial statistics, quantitative fluorescence microscopy, super-resolution microscopy

1 Introduction

The characterization of molecular interactions is a major challenge in quantitative microscopy. It is usually addressed in living cells by fluorescently labeling the interaction partners with spectrally distinct fluorophores and simultaneously imaging them. Typically, potential protein-protein inter-

1. ∗ F. Lavancier and T. Pécot contributed equally to this work.

actions inside the cell are determined by the degree of co-localization at the resolution limit of the microscope, that is the proportion of interacting proteins co-detected at the same location or in very close proximity [1, 2]. It often corresponds to co-compartmentalization, implying that two or more molecules bind to the same structure or domain in the cell. Accordingly, co-localization is known to be a critical step in the analysis of molecular interactions, but, for the time being, there is no definitive solution to this problem. Moreover, colocalization also needs to be correctly addressed to manage and objectively quantify the increasing amount of 2D-3D+time data, particularly with the emergence of super-resolution microscopy such as PALM (photoactivated localization microscopy), dSTORM (direct stochastic optical reconstruction microscopy) or SIM (Structured Illumination Microscopy).

In the literature, two distinct categories of co-localization approaches are generally considered [2, 3] : intensity-based methods [1, 4] and object-based methods [5, 6, 7] (see Section S1 in SI for a brief presentation). A commonly-used intensity-based technique is the Pearson's correlation method which returns values between -1 and +1 as a measure of the similarity of channels ; a positive correlation value corresponds to co-localization. This method is quite popular and is known to be superior to the Manders' method [8]. Nevertheless, the Pearson's method is sensitive to high intensity backgrounds and provides errors if the signal-to-noise ratios or the scales of the channels are not similar. Moreover, the choice of the threshold (typically 0.2) is arbitrary while it should be image dependent, taking into account the amount of available information (including the signal-to-noise ratio and the cell size) [8, 9]. As an alternative, object-based methods have been recently investigated [5, 6]. The detected spots (or regions) in cells are reduced to points (their centres) and the interactions between the points of the two channels are analyzed by spatial point processes statistics. This technique is particularly well-adapted for PALM and dSTORM images and when the objects of interest can be fairly assimilable to points (typically small balls). However, it is not suitable in presence of large or anisotropic shaped objects (*e.g.* microtubules, actin filaments), in which case the reduction of each object to a single point constitutes a dramatic loss of information. Finally, it is worth noting that all these methods cannot accurately compute colocalization between a pair of diffraction-limited (*e.g.* wide-field) and super-resolved images (*e.g.* PALM).

In this paper, we present a unique colocalization method (GcoPS) able to handle a large variety of scenarios in fluorescence microscopy. Given a pair of 2D/3D segmented images (or binary spot detection images) for each channel, GcoPS, only controlled by a p-value, tests whether the Pearson correlation between the two binary images is significantly positive. Unlike previous methods, GcoPS is robust to noise, fast (about two milliseconds per image pair in 2D and a few seconds in 3D with no dependence on the number of objects per image) mainly because it requires no simulation unlike [4] and versatile since it efficiently handles 2D and 3D images, variable SNRs and any kind of cell shapes and sizes. Also, for the first time, an algorithm is able to evaluate the colocalization between large regions and small dots, *e.g.* TIRF-PALM experiments, (see Fig. 2) and to interpret significant negative colocalization (or anti-colocalization). It provides results that are less sensitive to the presence of spurious isolated molecules, which may falsely influence point-based methods. The following section describes the underlying theory.

2 Mathematical formulation of the method

GcoPS applies a colocalization test on a 2D-3D binary image pair as explained in this section. The two binary images are obtained by segmenting the input fluorescence images. This splits the set of pixels into a background set and a foreground set, both of which should be non-empty. Any

segmentation algorithm can be used for this step if it provides a biologically reasonable segmentation of the tagged molecules. In the *Experimental Results* section, we show that GcoPS is fairly robust to the choice of segmentation algorithm parameters.

In what follows, we consider the binary images as realizations of random sets observed through a pixelated image. Formally, let Γ_1 and Γ_2 be two random sets in $\mathbb{R}^d$, where d stands for the dimension (typically $d = 2$ or $d = 3$ in our examples), and Ω be a pixelated image (that is the intersection of a lattice in $\mathbb{R}^d$ with a finite domain). In our mathematical setting, the binary images consist in the observation of $\Gamma_1 \cap \Omega$ and $\Gamma_2 \cap \Omega$. This probabilistic point of view is in agreement with the fact that the observations are very variable (so random) from an experiment to another, in particular the number of segmented objects, their positions, their sizes and their shapes are not constant.

Let us consider, for a generic given point $o \in \mathbb{R}^d$, the probabilities

$$p_1 = P(o \in \Gamma_1),\ p_2 = P(o \in \Gamma_2),\ p_{12} = P(o \in \Gamma_1 \cap \Gamma_2). \tag{1}$$

These quantities can be estimated from the observation of $\Gamma_1 \cap \Omega$ and $\Gamma_2 \cap \Omega$ by

$$\begin{aligned} \hat{p}_1 &= |\Omega|^{-1} \sum_{x \in \Omega} \mathbf{1}_{\Gamma_1}(x), \quad \hat{p}_2 = |\Omega|^{-1} \sum_{x \in \Omega} \mathbf{1}_{\Gamma_2}(x), \\ \hat{p}_{12} &= |\Omega|^{-1} \sum_{x \in \Omega} \mathbf{1}_{\Gamma_1}(x) \mathbf{1}_{\Gamma_2}(x). \end{aligned} \tag{2}$$

Here $|\Omega|$ denotes the number of pixels (or voxels) in Ω, and $\mathbf{1}_{\Gamma_1}(x)$ equals to 1 if $x \in \Gamma_1$ and to 0 otherwise (and similarly for $\mathbf{1}_{\Gamma_2}(x)$). These estimators are unbiased, *i.e.* $\mathbb{E}(\hat{p}_1) = p_1$ and so on, whenever Γ_1 and Γ_2 are stationary random sets, which means that their probabilistic law is invariant by translation.

If Γ_1 and Γ_2 are two independent sets (the case of no colocalization), we have $p_{12} = p_1 p_2$. A natural empirical measure of the departure from independence between Γ_1 and Γ_2 is therefore

$$D = \hat{p}_{12} - \hat{p}_1 \hat{p}_2. \tag{3}$$

Note that $\hat{p}_{12}$ is exactly the area (or volume if $d = 3$) of the intersection between Γ_1 and Γ_2 in Ω (up to $|\Omega|^{-1}$), so that D is the difference between the observed intersection area and the expected intersection area if Γ_1 and Γ_2 were independent. To make the connection with a more familiar interpretation, note that once properly normalized, D just becomes the empirical Pearson correlation between the two binary images $\Gamma_1 \cap \Omega$ and $\Gamma_2 \cap \Omega$:

$$-1 \ \leq \ \hat{\rho} = \frac{\hat{p}_{12} - \hat{p}_1 \hat{p}_2}{\sqrt{\hat{p}_1(1-\hat{p}_1)\hat{p}_2(1-\hat{p}_2)}} \ \leq \ 1. \tag{4}$$

If Γ_1 and Γ_2 are independent sets, then one expects $\hat{\rho}$ to be close to zero. In order to build a formal testing procedure, it is necessary to get the probability distribution of $\hat{\rho}$ (or D), or at least an approximation of the latter, as investigated in the following. It is worth emphasizing that the Costes method [4] follows the exact same procedure, except that the focus is on the PCC (Pearson Correlation Coefficient) between the intensity images and not the binary images. Then the probability distribution of the PCC is approximated in the Costes method by simulation, where a crucial tuning parameter (namely the size of the pixels' blocks, see Section S1 in SI) has to be chosen. In contrast, as detailed below, our method does not need any simulation procedure to approximate the probability distribution of $\hat{\rho}$ but exploits a closed-form formula.

Assume that Γ_1 and Γ_2 are stationary random sets and that Γ_1 is independent of Γ_2. Denote by C_1 and C_2 their auto-covariance functions, respectively. Specifically, for $r \in \mathbb{R}^d$, denoting $\Gamma_1 - r$ the translation of Γ_1 by the vector r,

$$C_1(r) = P(o \in \Gamma_1 \cap (\Gamma_1 - r)) - p_1^2, \tag{5}$$

where $P(o \in \Gamma_1 \cap (\Gamma_1 - r))$ is the probability that two points separated by r belong to Γ_1. In particular $C_1(0) = p_1(1-p_1)$. The formulas for C_2 are the same where Γ_1 is replaced by Γ_2, and p_1 by p_2. The expectation of D is zero ($\mathbb{E}(D) = 0$) and its variance V_D is asymptotically (*i.e.* as $|\Omega|$ tends to infinity) given by (see proof in Section S2 in SI) :

$$V_D \;\approx\; |\Omega|^{-2} \sum_{x\in\Omega} \sum_{y\in\Omega} C_1(x-y) C_2(x-y). \tag{6}$$

Under some assumptions on Γ_1 and Γ_2, then it can be proved [10] that $D/V_D^{1/2} \to \mathcal{N}(0,1)$ as $|\Omega|$ tends to infinity, where $\to$ stands for the convergence in distribution. For instance, this convergence holds true if Γ_1 and Γ_2 are m-dependent [11] for some $m > 0$. Recall that Γ_1 is m-dependent if the events $\mathbf{1}_{x\in\Gamma_1}$ and $\mathbf{1}_{y\in\Gamma_1}$ are independent whenever x and y are separated by a distance greater than m. This assumption implies that C_1 and C_2 are finite-range and so summable.

In practice, we approximate the variance V_D as follows

$$\hat{V}_D = |\Omega|^{-2} \sum_{x\in\Omega} \sum_{y\in\Omega} \hat{C}_1(x-y) \hat{C}_2(x-y), \tag{7}$$

where $\hat{C}_1$ and $\hat{C}_2$ simply correspond to the empirical auto-covariance functions of the binary images $\Gamma_1 \cap \Omega$ and $\Gamma_2 \cap \Omega$. For instance for $\Gamma_1 \cap \Omega$, an expression of $\hat{C}_1(r)$ can be

$$\hat{C}_1(r) = |\Omega|^{-1} \sum_{x\in\Omega} \sum_{y\in\Omega} \mathbf{1}_{\Gamma_1}(x) \mathbf{1}_{\Gamma_1}(y) \mathbf{1}_{x-y=r} - \hat{p}_1^2. \tag{8}$$

We finally obtain that if Γ_1 and Γ_2 are m-dependent stationary random sets and if Γ_1 is independent of Γ_2, then

$$\frac{D}{\sqrt{\hat{V}_D}} \to \mathcal{N}(0,1) \tag{9}$$

as $|\Omega|$ tends to infinity. In practice $|\Omega|$ represents the number of pixels and is typically large.

The testing procedure for independence between Γ_1 and Γ_2 is now straightforward. The test statistic, that we alternatively name *colocalization score*, is

$$T := \frac{D}{\sqrt{\hat{V}_D}} = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)\hat{p}_2(1-\hat{p}_2)}{\hat{V}_D}}\; \hat{\rho}. \tag{10}$$

As shown by these two expressions, we can either view T as a metric based on the intersection area of the two binary images, or as a normalized correlation coefficient between the two binary images. From Eq. [9], the null hypothesis of independence is rejected at the asymptotic level $\alpha \in (0,1)$ if $|T| > q(\alpha/2)$ where $q(\alpha/2)$ denotes the upper $\alpha/2$-quantile of the standard normal distribution. The corresponding p-value is

$$p\text{-value} = 2(1 - \Phi(|T|)), \tag{11}$$

where Φ denotes the cumulative distribution function of the standard normal distribution. This bilateral test can be modified into a unilateral test if the focus for the alternative hypothesis is

more specifically colocalization, *i.e.* attraction between Γ_1 and Γ_2, rather than dependence (in a general sense) between Γ_1 and Γ_2. In this case the null hypothesis becomes anti-colocalization, *i.e.* repulsion between Γ_1 and Γ_2. Then a more powerful procedure at the asymptotic level α consists in rejecting the null when $T > q(\alpha)$, which corresponds to p-value $= 1 - \Phi(T)$. Similarly, if the focus is the detection of anti-colocalization, one can apply p-value $= \Phi(T)$ for a more powerful testing procedure in comparison with Eq. [11].

3 Experimental results

3.1 Evaluations on synthetic data sets

We validated the performance of the GcoPS method on 2D and 3D data sets in adverse and noisy situations and compared it to the competitive Costes [4] and Lagache [6] methods. We refer to Section S1 in SI for a brief description of these procedures. We show here that GcoPS unequivocally outperforms the best competitive methods on dual-color images for different degrees of colocalization.

3.1.1 Detection of colocalization on simulated 2D images

We evaluated the sensitivity of the methods on synthetic images generated by the simulator described in [6]. This simulator consists in a first step to generate randomly distributed particles (say the red channel), and in a second step to simulate a proportion of green particles nearby red particles (this proportion is referred to as the *level of colocalization* in the sequel) while the rest of green particles are drawn randomly and independently. Three scenarios were considered (see Fig. 1 and Fig. S1) : i) without noise (Fig. 1), ii) with noise and iii) with noise and a spatial shift of three pixels between colocalized particles (a 3 pixels shift is more than enough to account for a spatial shift due to the experimental device). In each of these scenarios, the two channels were simulated with a level of colocalization of 0% (no colocalization), 2.5% and 5%. In each situation, 1000 pairs of images were generated to evaluate the sensitivity of the three tested methods.

Figure 1 compares the results of GcoPS, the Lagache method and the Costes method for the first scenario (no noise, no shift), by reporting the proportion of p-values lower than 0.05. Recall that for a perfect testing procedure, these proportions should be equal to 5% in absence of colocalization while they should be as close as possible to 1 in presence of colocalization. The results in Fig. 1 reveal that the Lagache method is not sufficiently sensitive (less than 35% of images have a p-value inferior to 0.05 for a 5% colocalization level against more than 90% for GcoPS). The Costes method is in turn too sensitive in absence of colocalization (about 18% of images without colocalization have a p-value lower than 0.05), which leads to too many false positive decisions. Fig. S1 confirms this conclusion by reporting, for all scenarios, the proportion of p-values lower than 0.05, along with the empirical distribution function (edf) of the p-values [2]. A perfect testing procedure would result in an edf equal to the first diagonal in absence of colocalization and in an edf which is uniformly equal to 1 in presence of colocalization. Fig. S1 demonstrates that GcoPS gives satisfying results in all situations without colocalization, although it seems a bit too sensitive in presence of noise and shift, and is more sensitive than the other methods in presence of colocalization (twice more images have a p-value lower than 0.05 for a 5% colocalization level in presence of noise and/or shift, see Figs. S1b-c).

2. This function counts, for any $x \in [0, 1]$, the proportion of p-values lower than x

3.1.2 Sensitivity to image segmentation and deconvolution

To evaluate the influence of the segmentation on the different methods, we considered the same simulated images with noise and shift than in Fig. S1c but with four different thresholds (see Fig. S2a), and then applied the same methods as described in the previous section. We got similar binarized images by applying a thresholding operator to a set of increasing blurred images. Given a unique thresholding value, we obtained large connected components for highly blurred images and objects with small areas if the image is slightly blurred. Consequently, we jointly demonstrate here the robustness of GcoPS to either imprecise image segmentation or image deconvolution. As a reference, we also include the segmentation obtained with the ATLAS spot detection method [12], which is the segmentation method used in all experiments presented in the manuscript. Note that the Costes method gave different results than in Fig. S1 as the input corresponds to the segmented particles here and not the intensity as in Fig. S1. Figure S2 shows that this method is not much affected by the pre-processing and is more sensitive than GcoPS when the level of colocalization is 2.5% and 5% (see Figs. S2c-d). However, the Costes method clearly leads to too many false positive decisions when there is no colocalization (see Fig. S2b). The Lagache method is clearly affected by over-segmentation ($\tau = \{15; 20\}$), in which case it completely fails to detect colocalization (see Fig. S2d). GcoPS also gives less satisfying results when the segmentation is not well processed, but the results are overall still satisfying. In particular GcoPS is very robust to pre-processing for images without colocalization, which is a safe guaranty against false positive decisions.

3.1.3 Sensitivity to the shape and scale of objects

To evaluate the sensitivity to the shape and the size (or scale) of objects, we simulated Gaussian level sets with a correlation between the two channels equal to 0 (*i.e.* no colocalization), 0.1 (slight colocalization) and 0.3 (stronger colocalization) approximately. This method of simulation is detailed in Section S3 in SI. It allows to generate objects that exhibit non-regular shapes (different from balls), with a typical scale that can be easily controlled. When the objects are small, the images are quite similar to the images generated in Fig. S1, or in other words the objects are fairly assimilable to small balls, and the methods show similar efficiency as described above. In presence of moderate size and non regular objects (see Fig. S3a), the Costes method is once again too sensitive for colocalization while the Lagache method is not sufficiently sensitive, especially when the correlation between the two channels is 0.1. As shown in Fig. S3a, the performances of GcoPS are not disrupted by the shape of objects. In presence of larger objects as depicted in Fig. S3b, the conclusions are similar. Finally, in the extreme case of very large objects (see Fig. S3c), the Lagache method, that reduces each object to a single point, completely fails to detect colocalization, which is easily explained by the dramatic loss of information induced by this reduction. In this case, the Costes method is far too sensitive (about 25% of false positive decisions in absence of colocalization), while GcoPS still performs well, though being a bit too sensitive when there is no colocalization. The robustness of GcoPS in presence of large objects demonstrated in Fig. S3b-c shows that this method can be applied in small windows in the image, where the scaling makes objects larger, hence opening the possibility to effectively localize the detection of colocalization.

3.1.4 Sensitivity in presence of a different optical resolution in each channel

To generate channels with a different optical resolution, we simulated Gaussian level sets where the scale parameter ruling the typical size of objects is different in each channel (see Section S3 in SI for details). The correlation between the two channels is as in the previous section $\rho = 0$,

$\rho = 0.1$ or $\rho = 0.3$. Two situations are summarized in Fig. S4, depending on whether the difference in optical resolution is moderate (Fig. S4a) or strong (Fig. S4b and Fig. 2). These results show that the object-based method of Lagache is clearly out-performed by the other methods in presence of a difference of optical resolution in each channel. The Costes method and GcoPS behave similarly well in the setting of Fig. S4a, while GcoPS exhibits better efficiency in the case of a strong difference of optical resolution (Fig. S4b and Fig. 2). This proves that GcoPS is able to process efficiently images for which a different microscopy technique is used for each channel.

3.1.5 Evaluation on simulated 3D images

We also performed simulations of 3D objects using Gaussian level sets (see Section S3 in SI) to generate channels with a correlation equal to 0, 0.1 and 0.3. The Lagache method shows very unsatisfying results when there is no colocalization (see Fig. S5). The same figure reveals that the Costes method is even more sensitive to colocalization in 3D than it is in 2D, leading to too many false positive conclusions when channels are independent. The results obtained in Fig. S5 with GcoPS on 3D images are in line with the results obtained with 2D images and demonstrate the better overall performances of this method. Note finally that we also performed complementary simulations in 3D, not displayed here, to assess the robustness of GcoPS against shape anisotropy (*e.g.* elongated shapes in 3D) and/or a low density of particles. The results demonstrate that GcoPS is capable to reliably handling a wide range of shapes.

3.2 Application to real 3D super-resolution imaging

We show the potential of GcoPS on biological samples observed by dSTORM [13] and 3D-TIRFM [14].

3.2.1 Colocalization of BDNF and vGlut proteins in dSTORM

First, we evaluated the colocalization between BDNF (brain-derived neurotrophic factor) proteins and vGlut, a vesicle marker for presynapses on an image acquired with direct stochastic optical reconstruction microscopy (dSTORM, see Fig. 3) [13]. While identifying BDNF particles is straightforward with a spot detector [12], the segmentation of vGlut is more difficult as it does not correspond to a regular shape. Consequently, we performed three different segmentations by thresholding the image with three different thresholds (see Figs. S6a-c). The p-value obtained with GcoPS is extremely low (p-value $= 0$) for the three segmentations of vGlut, which unequivocally proves colocalization between BDNF proteins and vGlut, a conclusion consistent with previously published results [13]. The Lagache method [6] needs a segmentation of vGlut with a high threshold (see Fig. S6c) to obtain a low p-value (p-value $= 0.005$). Segmentations with low thresholds (see Figs. S6a-b) give large objects, leading to a failure for this method (p-value $= 0.532$ and p-value $= 0.061$). The Costes method provided p-values close to 0 whenever the size of blocks for the permutation step is fixed to 2×2, 5×5 or 10×10 pixels^2 in the two channels. This result is in agreement with the conclusion of GcoPS. However the Costes method needed about 3 minutes to run while GcoPS only took 9 seconds. We then applied GcoPS to windows of size 50×50 pixels randomly located in the image to identify the regions of colocalization. The window size is here chosen sufficiently small to analyze local interactions in detail, while being sufficiently large with respect to the size of the objects to safely apply our testing procedure.

3.2.2 Colocalization of Langerin and Rab11a in 3D-TIRFM

In the second experiment, we performed a temporal acquisition with 3D-TIRFM [14] using wild-type RPE1 cells transfected with Langerin-mCherry and Rab11a-GFP (see Fig. 4a). We processed a dense colocalization map (see Fig. 4b) at time $t = 0$ by computing the colocalization score every 25 pixels at the medium plane on windows of size $50\times50\times10$ voxels. The spatial density of colocalization scores is then computed by applying a Gaussian kernel density method with global bandwidth selection [15] (equal to 5 in Fig. 4b). This representation is an alternative to the colocalization hits shown in Fig. 3 to identify the regions of colocalization. It has the advantage to discriminate between different levels of colocalization since a more colocalized region will have a higher colocalization score. Negative values of colocalization score correspond to anti-colocalization.

3.2.3 Spatiotemporal colocalization of Langerin-mCherry/Rab11a-GFP in 3D TIRFM

In the last experiment, we applied GcoPS on the Langerin-mCherry/Rab11a-GFP image sequence (see Fig. 4a), frame by frame. We then shifted frames between the two channels and applied GcoPS from -20 to +20 frames shift by considering Langerin-mCherry as the reference. The colocalization score is reported for each temporal shift in Fig. 4c. The slope of the colocalization level is steeper for a positive temporal shift than for a negative temporal shift, demonstrating that globally, Rab11a is visible before Langerin, which is consistent with Cinquin [16]. With some originality, this experiment also confirms the robustness of GcoPS to chromatic aberrations (or color shifts) due to improper matching of optical components or mechanical shifts, and assumed to be not well compensated. Actually, the two channels are expected to strongly co-localize. The comparison of image pairs with an increasing temporal shift can be seen as a way to mimic spatial shifts of channels. As said before, GcoPS confirms co-localization for any temporal shifts but the colocalization score decreases as the temporal shift increases.

4 Discussion

We developed GcoPS, an original, fast and robust approach to quantify interactions between molecules. It is a flexible method that joinlty handle image pairs with small spots and large regions with variable shape and size. This generalization not only makes the computation much faster (see Table S2) but results in a lower sensitivity to noise. The testing procedure is well-grounded in statistics, robust to different configurations and low signal-to-noise ratios. It is based on the Pearson's correlation and requires the setting of only one parameter corresponding to a p-value. In Fig. 5 and Tables S1-S2, we summarize the properties of GcoPS and of the two competing methods [4, 6] when the decision is based on p-values lower than 5%. In Fig. 5, the proportions correspond to the average of all situations shown in Figs. S1-S5, either in absence of colocalization (left part of Fig. 5) or in colocalized situations (right part of Fig. 5).

In a nutshell, the main drawbacks of the Costes method [4] is the CPU time (Table S2), that can be very large for 2D+t and 3D images, along with the detection of too many false positives (12% in average and up to 25%, instead of the expected 5%). The size of blocks is a critical parameter and influences the results of the permutation step in the procedure. The main weakness of the Lagache method [6] is the lack of sensitivity to colocalization. This appears clearly in Fig. 5 and becomes problematic in presence of large objects (Fig. S3), a situation where reducing each object to a point is unsuitable.

In turn, GcoPS does not suffer from the aforementioned drawbacks : i) it is fast ; ii) it well

controls the number of false positives ; iii) it is very sensitive to colocalization ; and iv) it is robust to segmentation, to the shape and the size of the objects and to a possible different optical resolution in each channel. A nice characteristic of GcoPS is the absence of tuning parameters given a pair of segmented images. As shown in the experiments, GcoPS is very robust to the choice of the segmentation algorithm and parameters. The reason is that the colocalization score in GcoPS is built from the area of the intersection of the two binary images, see Eq. [10], and this area is poorly affected by spurious points due to a bad segmentation algorithm. Note that Fig. S3 shows synthetic images with variable object sizes that can be interpreted as reconstructed images obtained with single-molecule localization microscopy techniques, such as PALM and STORM, relying on the cumulative spatial localization of fluorescently tagged markers. Instead of evaluating colocalization exclusively from the locations of detected molecules, GcoPS is able to exploit the localization uncertainty attached to the particle coordinates to make a robust decision.

For the practical implementation of GcoPS, it is important to determine the ratio between the size of the window with respect to the size and the number of objects in the window. This ratio should be "as large as possible" so that the GcoPS colocalization score is distributed as a standard Gaussian law as claimed in Eq. [9]. The control of the convergence to the Gaussian distribution, or rather the absence of control of this convergence, is a common issue in almost all testing procedures, including in [4, 6]. Nonetheless, where the objects are extremely large with respect to the size of the windows (see Fig. 3c), GcoPS still behaves satisfyingly, proving that it is very robust to the detection of colocalization in small sub-windows of a real image. For a safe decision, we finally recommend to use GcoPS in windows that are at least five times larger than the average size of the objects. This guarantees that a minimal fluorescence information is available to assess colocalization.

Acknowledgment

We thank Jean Salamero, Markus Sauer, Soeren Doose, Sarah Aufmolk, Perrine Paul-Gilloteaux and Fabrice Cordelières for assistance with experiments and for helpful insight in the preparation of the manuscript.

Références

[1] E.M. Manders, F.J. Verbeek, and J.A. Aten. Measurement of co-localization of objects in dual-colour confocal images. *J Microscopy*, 169, Pt 3 :375–382, 1993.

[2] S. Bolte and F. Cordelieres. A guided tour into subcellular colocalization analysis in light microscopy. *J Microscopy*, 224, Pt 3 :213–232, 2006.

[3] K.W. Dunn, M.M. Kamocka, and J.H. McDonald. A practical guide to evaluating colocalization in biological microscopy. *Am J Physiol Cell Physiol*, 300(4) :C723–742, 2011.

[4] S.V. Costes, D. Daelemans, E.H. Cho, Z. Dobbin, G. Pavlakis, and S. Lockett. Automatic and quantitative measurement of protein-protein colocalization in live cells. *Biophysical J*, 86(6) :3993–4003, 2004.

[5] J.A. Helmuth, G. Paul, and I.F. Sbalzarini. Beyond co-localization : inferring spatial interactions between sub-cellular structures from microscopy images. *BMC Bioinformatics*, 11 :372, 2010.

[6] T. Lagache, N. Sauvonnet, L. Danglot, and J.-C. Olivo-Marin. Statistical analysis of molecule colocalization in bioimaging. *Cytometry Part A*, 87(6) :568–579, 2015.

[7] J.F. Gilles, M. Dos Santos, T. Boudier, S. Bolte, and N. Heck. DiAna, an ImageJ tool for object-based 3D co-localizatin and distance analysis. *Methods*, 115 :55–64, 2017.

[8] J. Adler and I. Parmryd. Quantifying colocalization by correlation : the Pearson correlation coefficient is superior to the Mander's overlap coefficient. *Cytometry part A*, 77(8) :733–742, 2010.

[9] A.L. Barlow, A. MacLeod, S. Noppen, J. Sanderson, and C.J. Guérin. Colocalization analysis in fluorescence micrographs : verification of a more accurate calculation of Pearson's correlation coefficient. *Micros Microanal*, 16(6) :710–724, 2010.

[10] S. Mase. Asymptotic properties of stereological estimators of volume fraction for stationary random sets. *J Applied Probability*, 19(1) :111–126, 1982.

[11] F. Gotze, L. Heinrich, and C. Hipp. m-dependent random fields with analytic cumulant generating function. *Scandinavian J Statistics*, 22(2) :183–195, 1995.

[12] A. Basset, J. Boulanger, J. Salamero, P. Bouthemy, and C. Kervrann. Adaptive spot detection with optimal scale selection in fluorescence microscopy images. *IEEE T Image Processing*, 24(11) :4512–4527, 2015.

[13] T. Andreska, S. Aufmkolk, M. Sauer, and R. Blum. High abundance of BDNF within glutamatergic presynapses of cultured hippocampal neurons. *Front Cell Neurosci*, 8 :1–15, 2014.

[14] J. Boulanger, C. Gueudry, D. Munch, B. Cinquin, P. Paul-Gilloteaux, S. Bardin, C. GuÃ©rin, F. Senger, L. Blanchoin, and J. Salamero. Fast high-resolution 3D total internal reflection fluorescence microscopy by incidence angle scanning and azimuthal averaging. *Proc Natl Acad Sci USA*, 111(48) :17164–17169, 2014.

[15] B.S. Silverman. *Monographs on Statistics and Applied Probability*. Chapman & Hall, London, 1986.

[16] B. Cinquin. *titre*. PhD thesis, University of Paris Diderot, 2011.

[17] C. A. Schneider, W. S. Rasband, and K. W. Eliceiri. NIH Image to ImageJ : 25 years of image analysis. *Nat Methods*, 9 :671–675, 2012.

Methods

Implementation of algorithms

We use JACoP plugin [2] on ImageJ [17] to apply the Costes method : i) in the randomisation step, $n = 1000$ replications are considered ; ii) we chose pixels' blocks with a size corresponding to the PSF for simulated images [4], and corresponding to the average size of the objects for segmented images. For the Lagache method, we use the colocalization studio of Icy (`http://icy.bioimageanalysis.org`). The CPU time of GcoPS is small when implemented on C++ and is slightly less optimal with the Icy plugin (based on Java), nevertheless keeping much faster than the two alternative methods (see Table S2). We refer to Section S4 in SI for a brief description of the implementation of GcoPS.

BDNF and vGlut proteins and dSTORM acquisition

We refer to [13] for the description of data shown in Fig. 3.

RPE1 cell preparation

For the set of experiments shown in Fig. 4a, wild type RPE1 cells were grown in Dulbecco's Modified Eagle Medium : Nutrient Mixture F-12 (DMEM/F12) supplemented with 10% (vol/vol) FCS, in 6 well plates. RPE1 cells were transiently transfected with plasmids coding for Langerin-YFP and Langerin-mCherry or Rab11a-GFP and Langerin-mCherry using the following protocol : 2 μg of each DNAs, completed to 100 μL with DMEM/F12 (FCS free) were incubated for 5 min at room temperature. 6 μL of X-tremeGENE 9 DNA Transfection Reagent (Roche) completed to 100 μL with DMEM/F12 (FCS free), were added to the mix and incubated for further 15 min at room temperature. The transfection mix was then added to RPE1 cells grown one day before and incubated further at 37^{o}C overnight. Cells were then spread onto fibronectin Cytoo chips (Cytoo Cell Architect) for 4h at 37^{o}C with F-12 (with 10% (vol/vol) FCS (feotal calf serum), 10 mM Hepes, 100 units/ml of penicillin and 100 μg/ml of Strep) before imaging. Cell adhesion on micropatterns both constrains the cells in terms of lateral movement and averages their size and shape (1100 μm2).

Dual color multi-angle TIRFM and 3D image acquisition

Live-cell imaging was performed using simultaneous dual color Total Internal Reflection Fluorescence (TIRF) microscopy. All imaging was performed in full conditioned medium at 37^{o}C and 5% CO_2 unless otherwise indicated. Simultaneous dual color TIRFM microscopy sequences were acquired on a Nikon TE2000 inverted microscope equipped with a x100 TIRF objective (NA=1.49), an azimuthal (spinning) TIRF module (Ilas2, Roper Scientifc), an image splitter (DV, Roper Scientific) installed in front of an EMCCD camera (Evolve, Photometrics) and a temperature controller (LIS). GFP and m-Cherry were excited with a 488 nm and a 561 nm laser, respectively (100mW). The system was driven by the Metamorph software (Molecular Devices). A range of angles corresponding to a set of penetration depths is defined for a given wavelength and optical index of the medium [14]. We performed simultaneous double-fluorescence image acquisition using RPE1 cells double transfected with Langerin-YFP and Langerin-mCherry or Rab11A-GFP and Langerin-mCherry. Image series corresponding to simultaneous two colors multi-angles TIRF image stacks were recorded at one stack of 12 angles every 360 ms during 14.76 s, with a 30 ms exposure time per frame. Three-dimensional reconstructions of the whole cells were performed on the first 300 nm in depth of the cells using a 30-nm axial pixel size (see [14]).

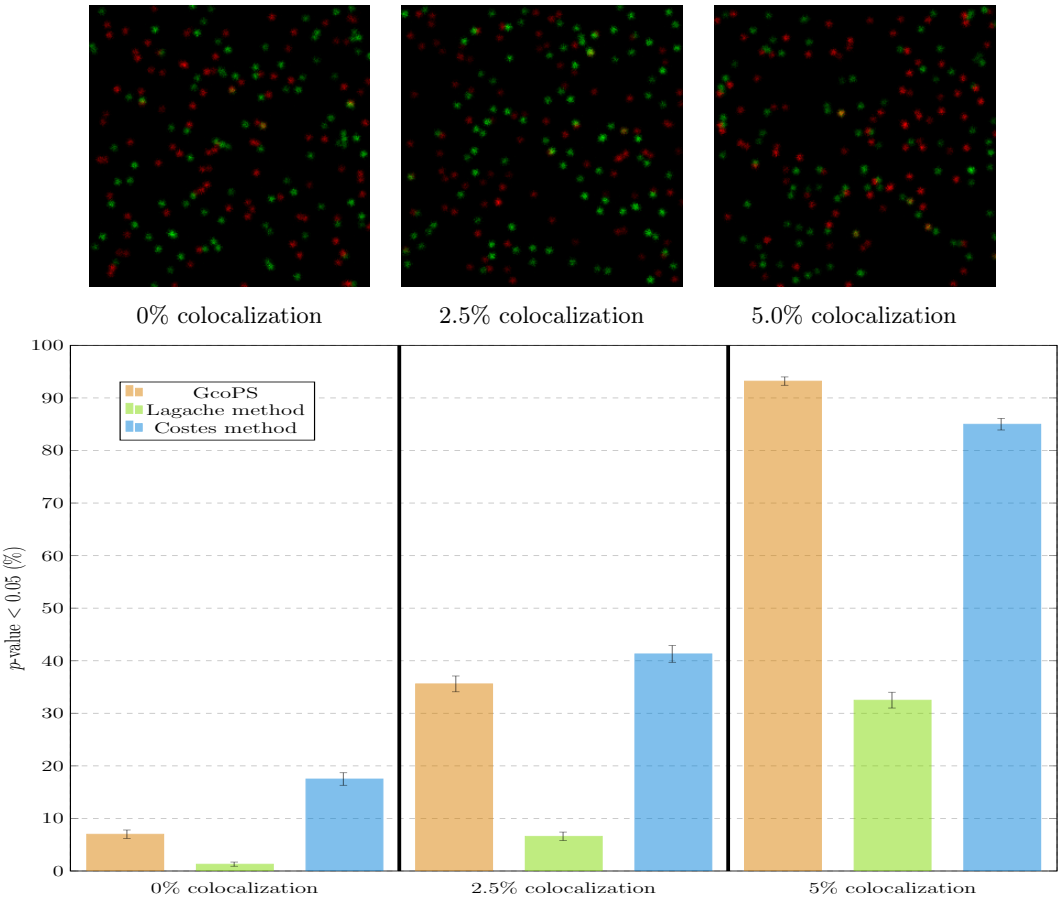


FIGURE 1 – **Detection of colocalization in noise-free images.** Proportion of p-values lower than 0.05 (bottom) obtained with GcoPS, the Lagache method and the Costes method over 1000 simulated noise-free images for 0%, 2.5% and 5.0% colocalization levels (left to right). The confidence intervals at the top of each bar represent one standard deviation over all 1000 simulations.

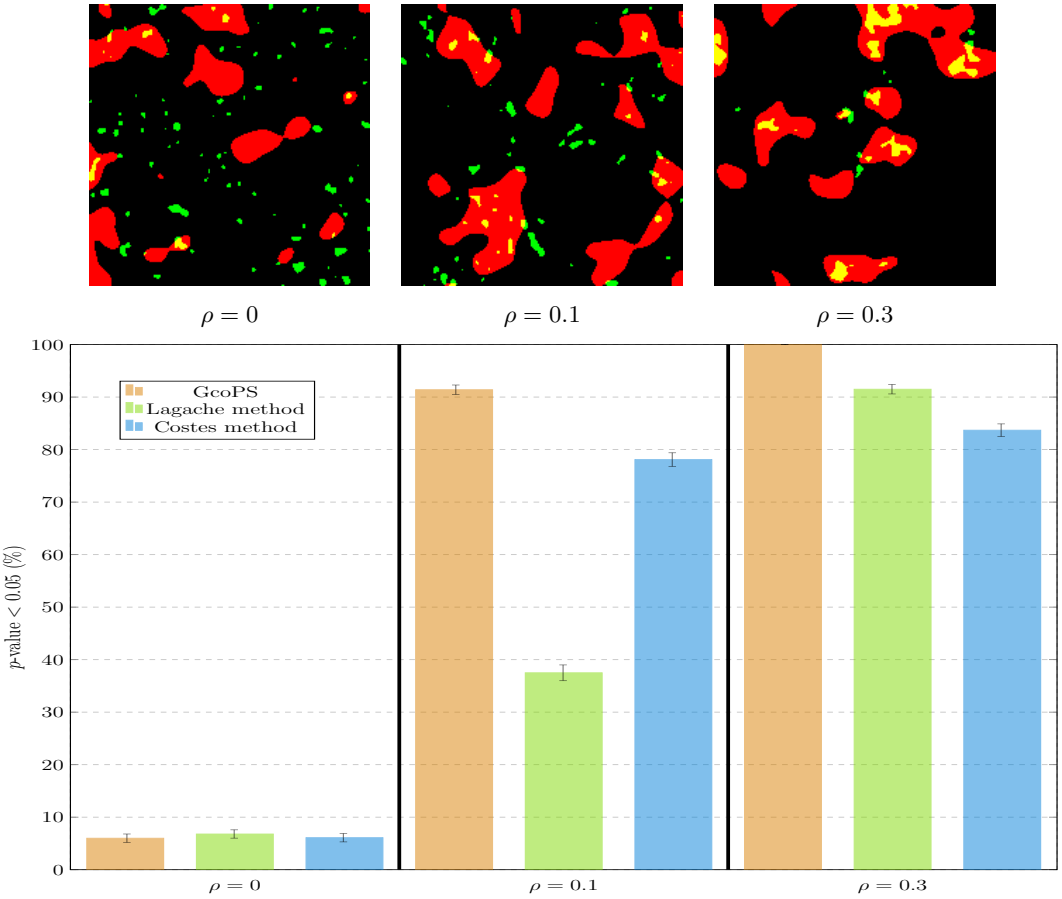


FIGURE 2 – **Sensitivity to the resolutions in each channel.** Proportion of p-values lower than 0.05 (bottom) obtained with GcoPS, the Lagache method and the Costes method over 1000 simulated images obtained via Gaussian level sets with a correlation equal to 0, 0.1 and 0.3 (left to right) between the two channels, where the resolution is very different in each channel.

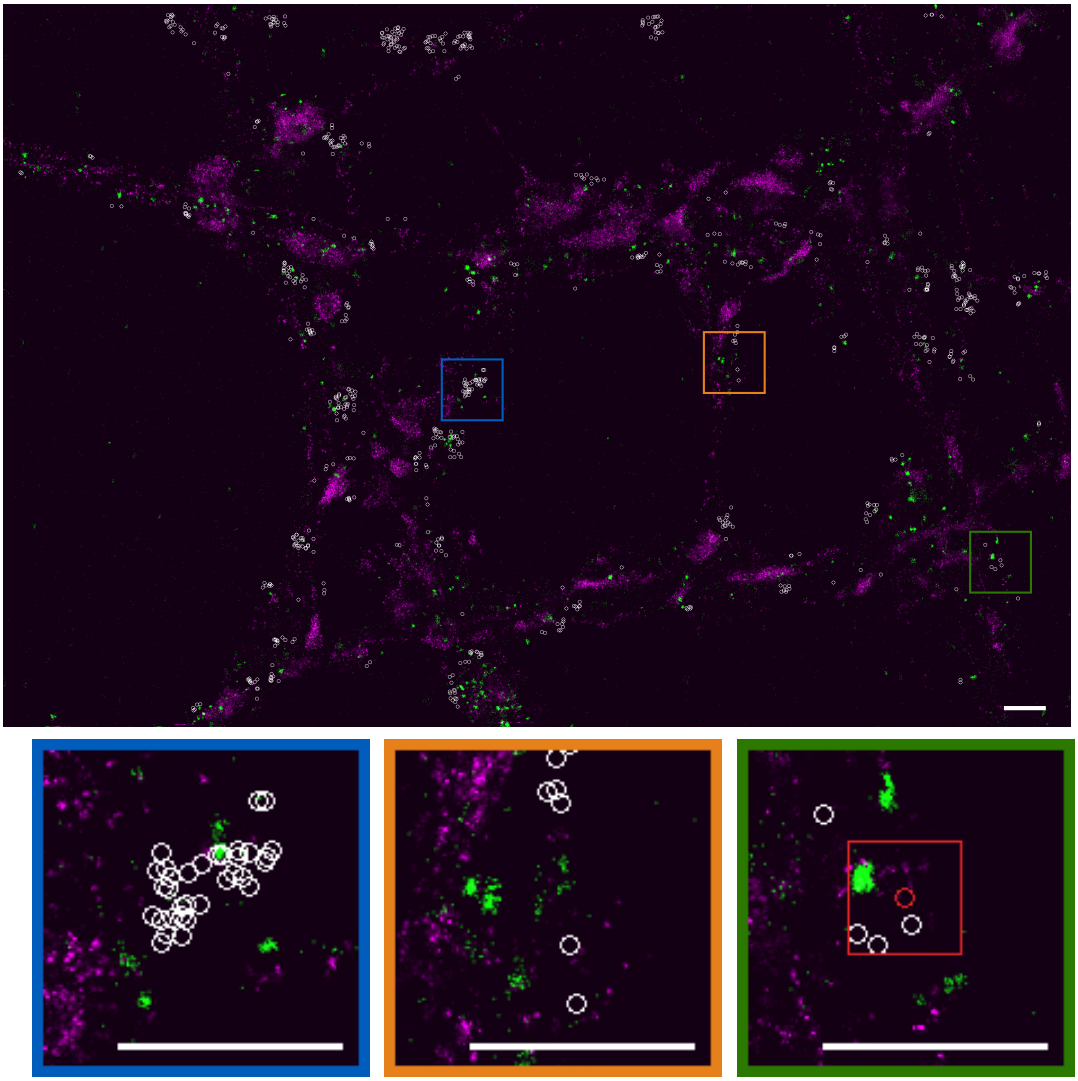


FIGURE 3 – **Colocalization of BDNF and vGlut proteins in dSTORM images.** DSTORM acquisition of cells from hippocampi of mice expressing BDNF proteins (green channel) and vGlut (purple channel) [13], with three zoomed-in regions displayed at the bottom. The colocalization regions identified by GcoPS are represented as white circles. The red rectangle represents the window used to find the hit shown as a red circle. The scale bars correspond to 1μm.

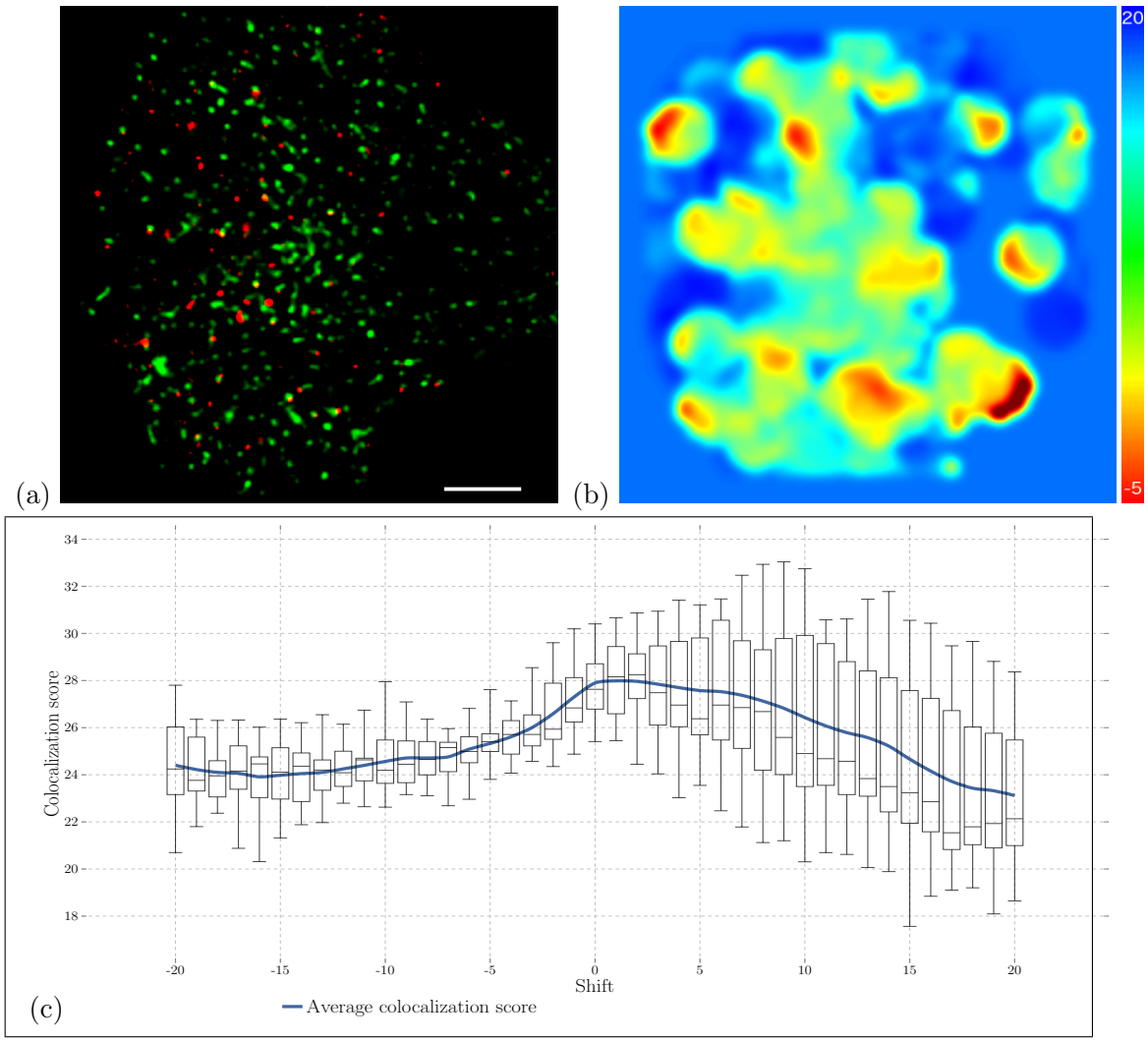


FIGURE 4 – **Spatiotemporal colocalization of Langerin and Rab11a proteins in multi-angles TIRFM images.** (a) Average intensity projection (along the z axis) of a 3D multi-angles TIRF acquisition [14] at time $t = 0$ showing a RPE1 cell expressing m-Cherry Langerin (red channel) and GFP Rab11 (green channel) on a crossbow shaped micropattern. The scale bar corresponds to 10 μm. (b) Heat map for continuous visualization of colocalization score associated to the left image at time $t = 0$. (c) Distribution of colocalization scores (box plots) and their averages (blue solid line) between pairs of images taken at time t and $t \pm \Delta t$, for a set of temporal shifts $\Delta t \in [-20, 20]$ and computed over the 60 image sequence ($t \in [0, 60]$) as displayed at the top at $t = 0$.

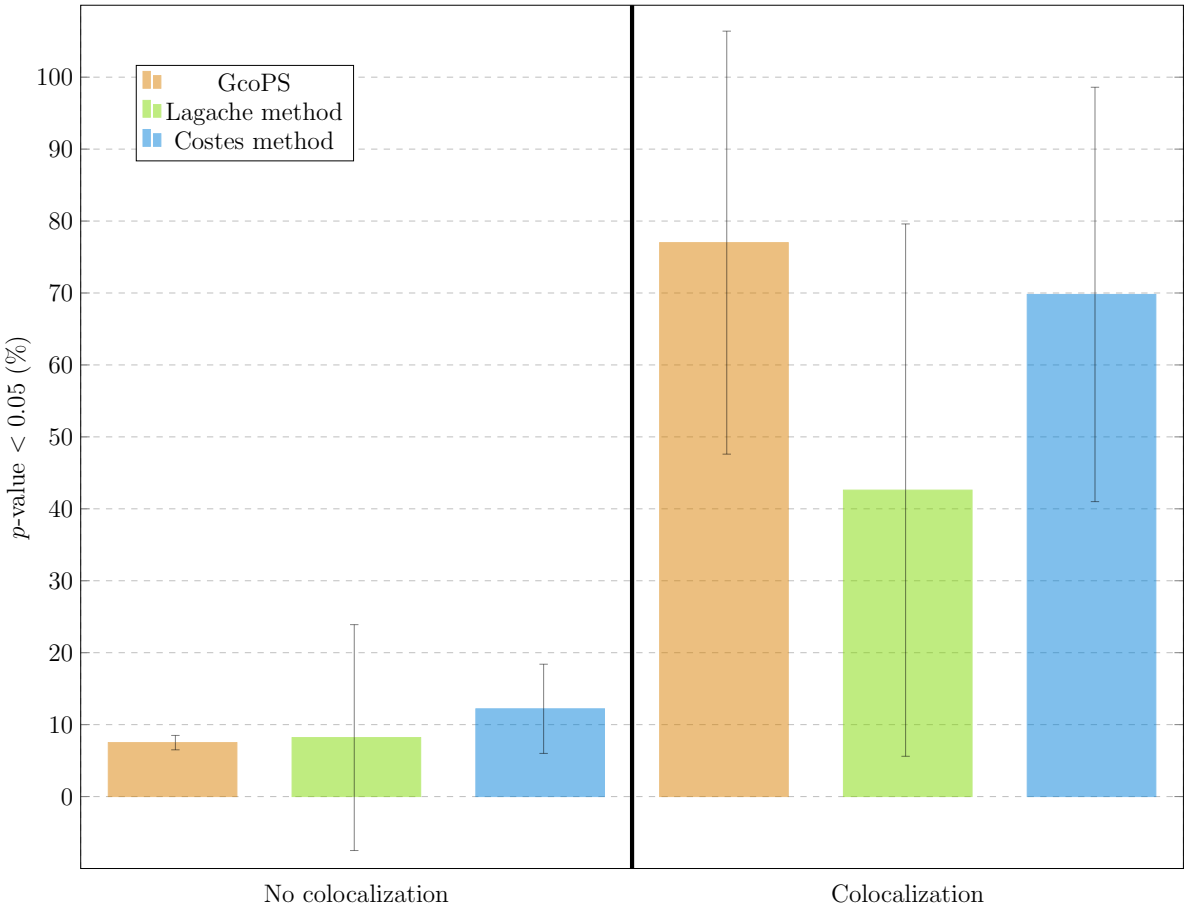


FIGURE 5 – **Proportion of colocalization detection.** Each bar represents the proportion of image-pairs detected as colocalized by GcoPS, the Lagache method and the Costes method, when the positive decision is based on a p-value lower than 5%. These proportions are the average proportions from all situations considered in Figures S1-S5, either when there is no colocalization (left part of the figure) or when there is colocalization (right part of the plot). The intervals at the top of each bar illustrates the variability of these proportions over all situations in Figures S1-S5. They correspond to one standard deviation of the associated proportions. In absence of colocalization, these proportions should be equal to 5%. In presence of colocalization, the proportions should be as close as possible to 1.

Supporting Information

S1. The colocalization methods

The most common metrics to analyze colocalization are the Pearson correlation coefficient (PCC) and the Manders coefficients, the expressions of which are recalled in the following. These coefficients are fast to compute and very popular. However, their outputs are strongly influenced by the image background, the presence of noise and/or shift, and the amount of available information (as the number of observed molecules in each channel). This makes impossible to define a clear and objective threshold for which there is actual colocalization in all situations. As a consequence, despite their popularity, these coefficients do not provide any decision rule. In practice, the users set an arbitrary threshold based on their experience.

To overcome this limitation, alternative methods have been introduced. We focus in our comparison study on the Costes method and the Lagache method, described below. The former is a way to assess the statistical significance of PCC, leading as an output to a p-value, and thus to a clear decision rule. The latter is an object-based method, that also provides a p-value. Typically, for both methods, colocalization is significantly detected if the p-value is less than 5%.

We now briefly describe the aforementioned methods, along with GcoPS. We denote by Ω the image domain, by $I_1(x)$ and $I_2(x)$ the intensities observed at site $x \in \Omega$ in channels 1 and 2, respectively, and by $\overline{I}_1$ and $\overline{I}_2$ the average intensities for channels 1 and 2, respectively.

Pearson correlation coefficient

$$PCC = \frac{\sum_{x\in\Omega}(I_1(x) - \overline{I}_1)(I_2(x) - \overline{I}_2)}{\sqrt{\sum_{x\in\Omega}(I_1(x) - \overline{I}_1)^2 \sum_{x\in\Omega}(I_2(x) - \overline{I}_2)^2}}.$$

Method parameters : pre-processing of the two images to alleviate the background effect, the possible presence of noise and/or shift between the two channels.

Manders coefficient

The two images are first segmented and we denote by $B_1(x)$ and $B_2(x)$ the values of the resulting binary images at site $x \in \Omega$. We agree that $B_1(x) = 1$ if x belongs to an object in the first image and $B_1(x) = 0$ otherwise, and similarly for $B_2(x)$. The Manders coefficients are defined by

$$MCC_1 = \frac{\sum_{x\in\Omega} I_1(x)B_2(x)}{\sum_{x\in\Omega} I_1(x)}, \qquad MCC_2 = \frac{\sum_{x\in\Omega} I_2(x)B_1(x)}{\sum_{x\in\Omega} I_2(x)}.$$

Method parameters : segmentation parameters, pre-processing of the two images to alleviate the background effect, the possible presence of noise and/or shift between the two channels.

Costes method

1. The PCC between the two channels is computed.
2. Image 1 is splitted into blocks of pixels (the size of these blocks corresponds by default to the PSF of the image in the JACoP plugin of ImageJ).
3. The blocks in Image 1 are randomly permuted and the PCC between the resulting image and Image 2 is computed.
4. The last two steps are repeated n times, leading to n values of PCC ($n = 1000$ by default in the JACoP plugin of ImageJ).
5. The proportion of PCCs in step 4 larger than the observed PCC in step 1 defines the p-value of the colocalization test.

Method parameters : pre-processing of the two images to alleviate the background effect, the possible presence of noise and/or shift between the two channels, size of the blocks, number of replications.

Lagache method

1. Each image is segmented into a set of objects and the mass center of each object is computed. This leads to two point patterns A_1 and A_2.
2. The Ripley's K function between A_1 and A_2, for several distances $r > 0$, is computed. This quantity, denoted by $K_{12}(r)$, counts the number of r-close pairs of points (one from A_1, one from A_2), up to a boundary correction term.
3. The p-value of the colocalization test is $\Phi(\max_r S(r))$ where $S(r)$ is a normalized version of $K_{12}(r)$ and Φ denotes the cumulative distribution function of the standard normal distribution.

Method parameters : segmentation parameters, range of values for $r \in \{r_{\min}, \ldots, r_{\max}\}$.

GcoPS

1. Each image is segmented into a set of objects, leading to two binary images.
2. The PCC between these two binary images is computed.
3. The p-value of the colocalization test is $\Phi(T)$, where T is a normalized version of the latter PCC (see Eq. [10] in Text).

Method parameters : segmentation parameters.

S2. Calculation of variance of D

We denote V_D the variance of D defined as

$$\begin{aligned}\mathrm{Var}(D) &= |\Omega|^{-2}\mathbb{E}\left(\sum_{x\in\Omega}\mathbf{1}_{\Gamma_1}(x)\mathbf{1}_{\Gamma_2}(x) - |\Omega|^{-1}\sum_{x\in\Omega}\mathbf{1}_{\Gamma_1}(x)\sum_{x\in\Omega}\mathbf{1}_{\Gamma_2}(x)\right)^2\\ &= V_1 + V_2 + V_3\end{aligned}$$

where

$$\begin{aligned}V_1 &= |\Omega|^{-2}\sum_{x\in\Omega}\sum_{y\in\Omega}C_1(x-y)C_2(x-y),\\ V_2 &= -2|\Omega|^{-3}\sum_{x\in\Omega}\left(\sum_{y\in\Omega}C_1(x-y)\right)\left(\sum_{y\in\Omega}C_2(x-y)\right),\\ V_3 &= |\Omega|^{-4}\sum_{x\in\Omega}\sum_{y\in\Omega}C_1(x-y)\sum_{x\in\Omega}\sum_{y\in\Omega}C_2(x-y).\end{aligned}$$

If C_1 and C_2 are summable auto-covariance functions, then $V_1 = O(|\Omega|^{-1})$, $V_2 = O(|\Omega|^{-2})$ and $V_3 = O(|\Omega|^{-2})$ as $|\Omega|$ tends to infinity, whereby $V_D \sim V_1$ as $|\Omega| \to \infty$.

S3. Simulation by Gaussian level sets

Let X, Y and ϵ be three independent Gaussian random fields in $\mathbb{R}^d$ with isotropic covariance function

$$C(r) = \sigma^2 e^{-r^2/\alpha^2},$$

where r denotes the radial distance between two points of the field, σ^2 is the variance and $\alpha > 0$ is referred to as the scale parameter. These parameters are denoted by σ_X^2, σ_Y^2, σ_ϵ^2 and α_X, α_Y, α_ϵ, for X, Y and ϵ respectively. Henceforth, for $\rho_0 \in [-1, 1]$, we set

$$\sigma_X^2 = \sigma_Y^2 =: \sigma_0^2 \quad \text{and} \quad \sigma_\epsilon^2 = \frac{\rho_0}{1-\rho_0}\sigma_0^2.$$

We define the random fields

$$U = X + \epsilon \quad \text{and} \quad V = Y + \epsilon.$$

Note that U and V are Gaussian random fields with common variance $\sigma^2 = \sigma_0^2/(1-\rho_0)$ and that U and V are correlated with correlation ρ_0.

We finally consider the random sets induced by the level sets of U and V, for $\tau_1 > 0$ and $\tau_2 > 0$,

$$\Gamma_1 = \mathbf{1}_{U>\tau_1\sigma} \quad \text{and} \quad \Gamma_2 = \mathbf{1}_{V>\tau_2\sigma}.$$

We easily get the following properties :

- The random set Γ_1, respectively Γ_2, has a coverage rate equal to $p_1 := E(\Gamma_1) = 1 - \Phi(\tau_1)$, respectively $p_2 := 1 - \Phi(\tau_2)$, where Φ denotes the cumulative distribution function of a standard normal law. Recall that this coverage rate represents the proportion of 1's, in average, generated by the binary field Γ_1 in a given domain.
- The random sets Γ_1 and Γ_2 are correlated with correlation

$$\rho = \left(\int_{\tau_1}^{\infty}\int_{\tau_2}^{\infty} f(u,v)\mathrm{d}u\mathrm{d}v - p_1p_2\right)/\sqrt{p_1(1-p_1)p_2(1-p_2)}$$

where f denotes the marginal probability density function of (U, V) that is the density of a bivariate centered Gaussian random variable with covariance matrix $\sigma^2 \begin{bmatrix} 1 & \rho_0 \\ \rho_0 & 1 \end{bmatrix}$.

In order to generate two correlated binary images containing random spots in a given domain, say Ω, we therefore simulate Γ_1 and Γ_2 in Ω. The input parameters are first the scale parameters α_X, α_Y and α_ϵ that rule the size of the spots (the larger the scale parameters, the larger the spots), second the thresholds τ_1 and τ_2 that rule the density of spots in Ω (see the expression of p_1 and p_2), third ρ_0 and σ_0 that along with the thresholds influence the correlation ρ between the two channels.

Given the input parameters, the simulation is straightforward. It basically amounts to simulate the Gaussian random fields X, Y and ϵ on Ω. We use at this step the `RandomFields` (SI-1) package of the free available software `R` (SI-2). Then the random fields U and V, and finally the random sets Γ_1 and Γ_2, are easily deduced.

In the simulations of Fig. S3, we used as input parameters $\sigma_0 = 1$, $\tau_1 = \tau_2 = 1$, resulting in a density of spots of $p_1 = p_2 \approx 16\%$, and $\rho_0 = 0, 0.2, 0.5$ leading to an actual correlation between the two channels of $\rho = 0, 0.1, 0.3$ approximately. The domain of simulation was $\Omega = [0, 250]^2$. As to the scale parameters (that have no impact on the values of p_1, p_2 and ρ), we chose $\alpha_X = \alpha_Y = \alpha_\epsilon = 8, 20, 50$ resulting in small, large or very large spots in Fig. S3. Concerning Fig. S4, the difference of optical resolution in the two images can be controlled by different scale parameters and/or thresholds parameters in the two channels. We set $\alpha_X = 5$, $\alpha_Y = \alpha_\epsilon = 10$, $\tau_1 = 1.5$, $\tau_2 = 1$ in Fig. S4a, and $\alpha_X = 5$, $\alpha_Y = \alpha_\epsilon = 20$, $\tau_1 = 2$, $\tau_2 = 1$ in Fig. S4b. The value of ρ_0 in these simulations has been tuned to result in the same final correlation between the channels as in Fig. S3, namely $\rho = 0, 0.1$ and 0.3 approximately. Finally, in Fig. S5, the domain of simulation is $\Omega = [0, 250]^2 \times [0, 60]$ and the input parameters are exactly the same as in Fig. S3 with the choice $\alpha_X = \alpha_Y = \alpha_\epsilon = 8$.

S4. Implementation of GcoPS

From a computational point of view, the cost reduces to the computation of T (see Eq. [10] in Text), where $\hat{p}_1$, $\hat{p}_2$, $\hat{p}_{12}$ and $\hat{\rho}$ are immediately obtained from the mean of the binary images $\Gamma_1 \cap \Omega$ and $\Gamma_2 \cap \Omega$ and their product. As to $\hat{V}_1$, its formula simplifies when Ω is an hypercube, say without loss of generality $\Omega = [1, n_1] \times \cdots \times [1, n_d]$:

$$\begin{aligned} \hat{V}_D &= \frac{1}{n_1^2 \dots n_d^2} \sum_{k_1=-n_1+1}^{n_1-1} \sum_{k_2=-n_2+1}^{n_2-1} \cdots \\ &\qquad \cdots \sum_{k_d=-n_d+1}^{n_d-1} \hat{C}_1(k)\hat{C}_2(k)\,(n_1 - |k_1|)\dots(n_d - |k_d|), \end{aligned}$$

where $k = (k_1, \dots, k_d)$ and the empirical covariance functions $\hat{C}_1$ and $\hat{C}_2$ are easily obtained by an FFT (Fast Fourier Transform) in each channel. Finally, there is no loss in practice to truncate the sums above since $\hat{C}_1(k)\hat{C}_2(k)$ becomes negligible for large values of k, and can be viewed as a nuisance noise. In the procedure GcoPS, we apply a truncation at $k_{\max}$ corresponding to the larger index such that both $\hat{C}_1(k)/\hat{C}_1(0) > 0.1$ and $\hat{C}_2(k)/\hat{C}_2(0) > 0.1$, in other words $k_{\max}$ is the maximal range of correlation in Γ_1 and Γ_2 beyond which the correlation is less than 0.1. This truncation does not deteriorate the estimation of V_1, but in turn speeds up significantly the computation.

References

(SI-1) Schlather M. et al. RandomFields : Simulation and Analysis of Random Fields. R package version 3.1.24.1 (2016) (`https://cran.r-project.org/package=RandomFields`).

(SI-2) R Core Team. R : A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria (2016) (`ttps://www.R-project.org`)

Fig. S1. Evaluation of GcoPS on 2D simulated images. Top right : proportion of p-values lower than 0.05 obtained with GcoPS, the Lagache method and the Costes method over 1000 simulated images without noise (**a**), with noise (**b**) and with noise and shift (**c**), for 0%, 2.5% and 5% colocalization levels. The confidence intervals at the top of each bar represent one standard deviation over all 1000 simulations. Bottom right : empirical distribution functions (edf) of p-values for the previous simulations. The edf counts, for any $x \in [0, 1]$ on the x-axis, the proportion of p-values less than x. Left : an example of simulated images.

Fig. S2. Sensitivity to segmentation and deconvolution. (**a**) Example of a simulated image with noise and shift and the corresponding segmentations with ATLAS and a threshold ranging from 15 to 30. (**b-d**) Proportion of p-values lower than 0.05 (bottom right) and empirical distribution functions of p-values obtained with (from top left to bottom left) GcoPS, the Lagache method and the Costes method over 1000 simulated images with noise and shift, segmented with ATLAS or a threshold ranging from 15 to 30, for 0% (**b**), 2.5% (**c**) and 5% (**d**) colocalization levels.

Fig. S3. Sensitivity to the shape and scale of objects. Proportion of p-values lower than 0.05 (top right) and empirical distribution functions (bottom right) of p-values obtained with GcoPS, the Lagache method and the Costes method over 1000 simulated images obtained via Gaussian level sets with a correlation equal to 0, 0.1 and 0.3 between the two channels, resulting in moderate size (**a**), large (**b**) and very large (**c**) non regularly shaped objects. An example of simulated images is shown at the left of each plot.

Fig. S4. Sensitivity to the resolutions in each channel. Proportion of p-values lower than 0.05 (top right) and empirical distribution functions (bottom right) of p-values obtained with GcoPS, the Lagache method and the Costes method over 1000 simulated images obtained via Gaussian level sets with a correlation equal to 0, 0.1 and 0.3 between the two channels, where the resolution is different (**a**) or very different (**b**) in each channel. An example of simulated images is shown at the left of each plot.

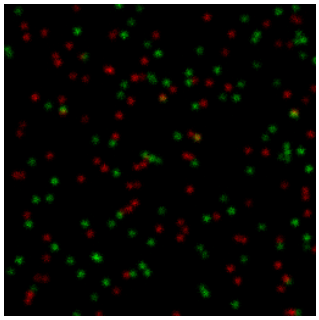
Fig. S5. Evaluation on 3D particles. Proportion of p-values lower than 0.05 (top right) and empirical distribution functions (bottom right) of p-values obtained with GcoPS, the Lagache method and the Costes method over 1000 simulated images obtained via 3D Gaussian level sets with a correlation equal to 0, 0.1 and 0.3 between the two channels. An example of simulated images is shown at the left of the plot.

Fig. S6. vGlut segmentations. Three different segmentations of purple channel of image shown in **Fig. 3** (see Text).

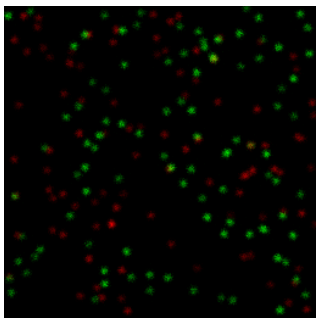
Table S1. Comparative summary. The numbers of stars, from 1 (weak) to 4 (good), stand as an informal score for each criterion, deduced from the evaluation results shown in Figs. S1-S5.

Table S2. Evaluation of computation time. Comparisons of computation time of the Costes method, the Lagache method and GcoPS, when applied to 2D, 2D+time and 3D images. As expected, the Costes procedure, which is not an object-based method, is not sensitive to the number of objects. However, it is by far the slowest method because of the randomisation step ($n = 1000$ replications by default). On the contrary, the Lagache method strongly depends on the number of objects and can therefore be quite slow if this number is large, which is typically the case in 3D (about 15 times slower then GcoPS in 2D if the number of objects is higher than 3500 objects. In contrast, GcoPS is very fast compared to the Costes procedure (about 8 times in 2D and 3D) whatever the situation is and is not sensitive to the number of objects. The CPU time of GcoPS is of course minimal when implemented on C++ and is slightly less optimal with the Icy plugin (based on Java), nevertheless keeping much faster than the two alternative methods.

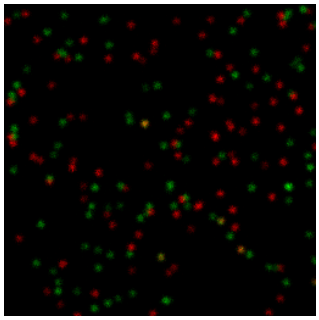
S1-a: No noise, no shift



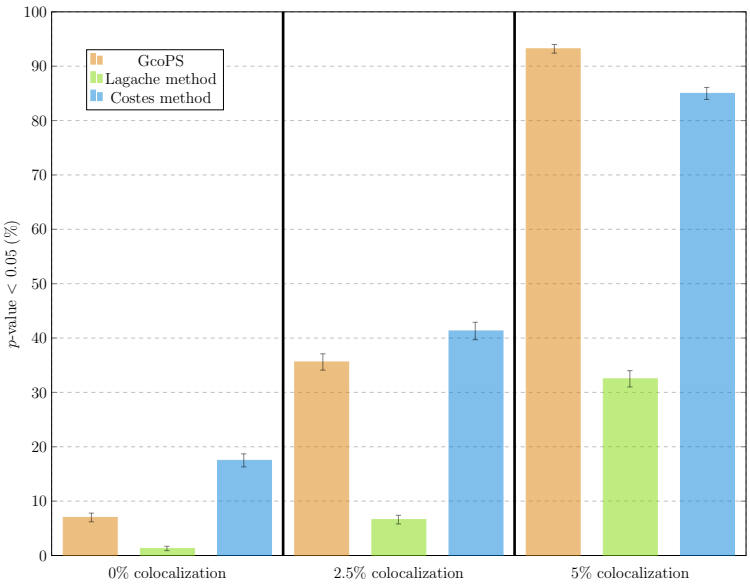
0% colocalization

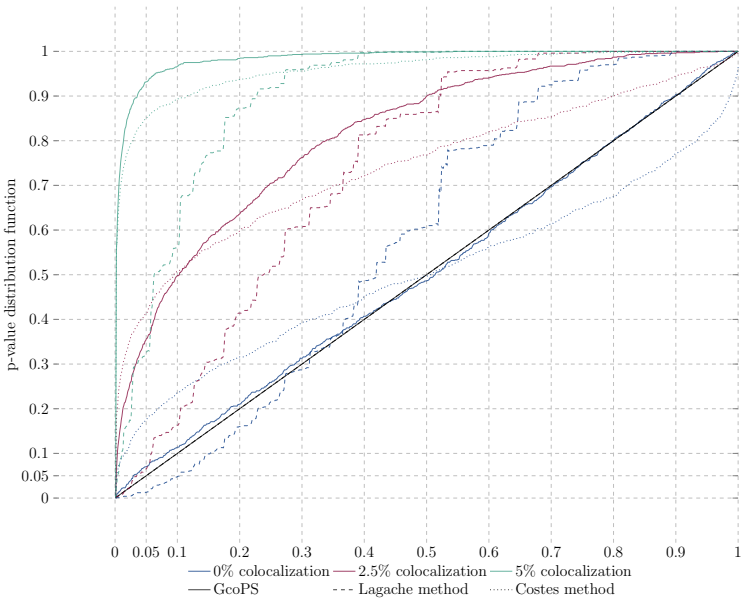


2.5% colocalization

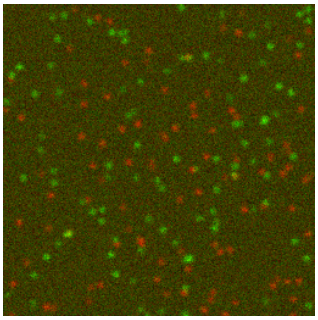


5% colocalization

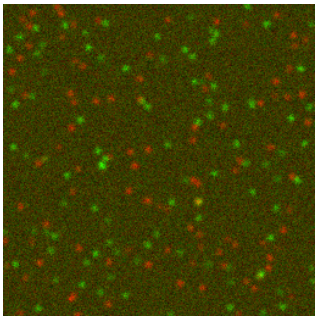




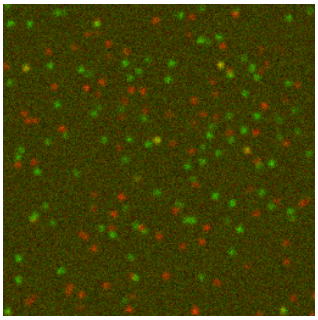
S1-b: Noise, no shift



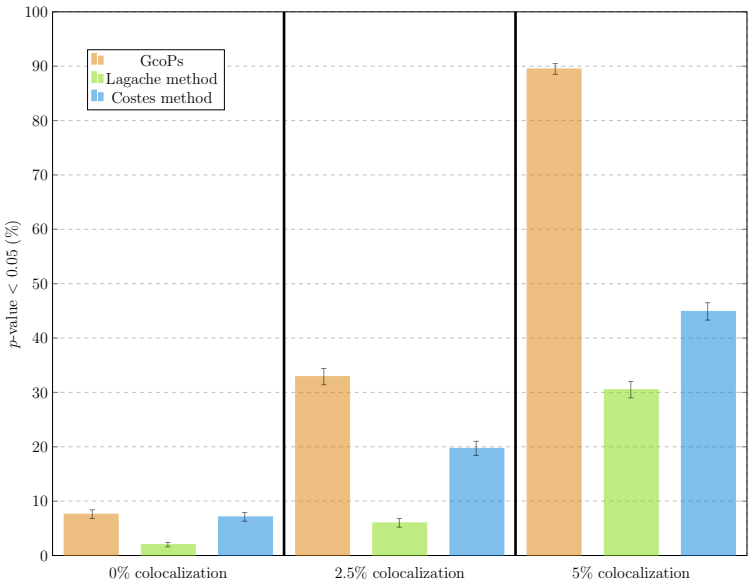
0% colocalization

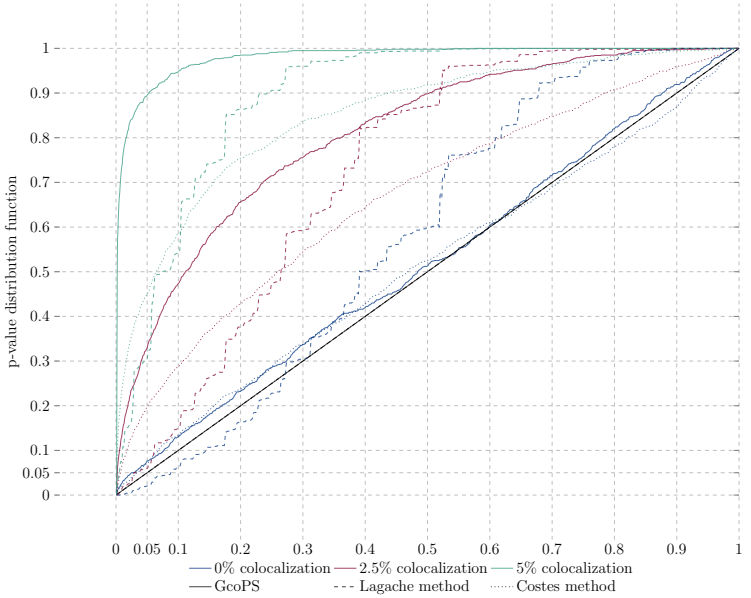


2.5% colocalization

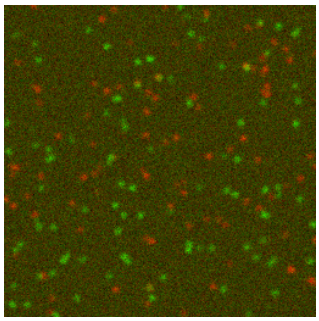


5% colocalization

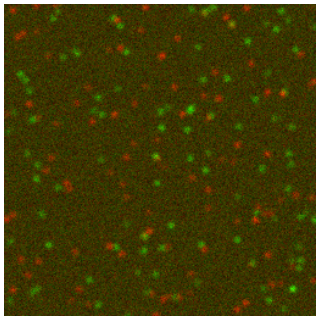




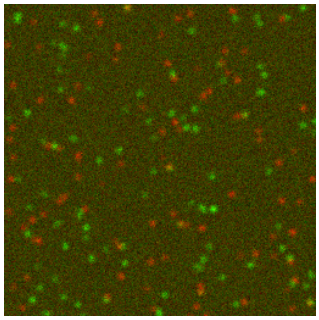
S1-c: Noise and shift



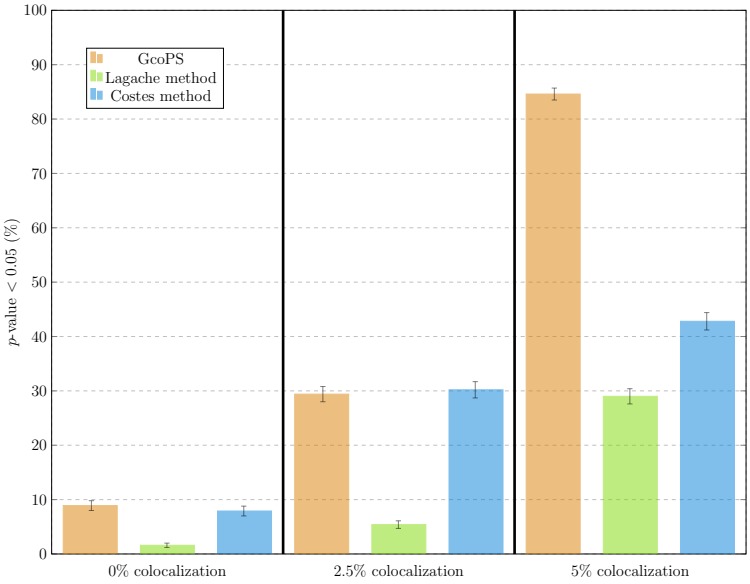
0% colocalization



2.5% colocalization



5% colocalization



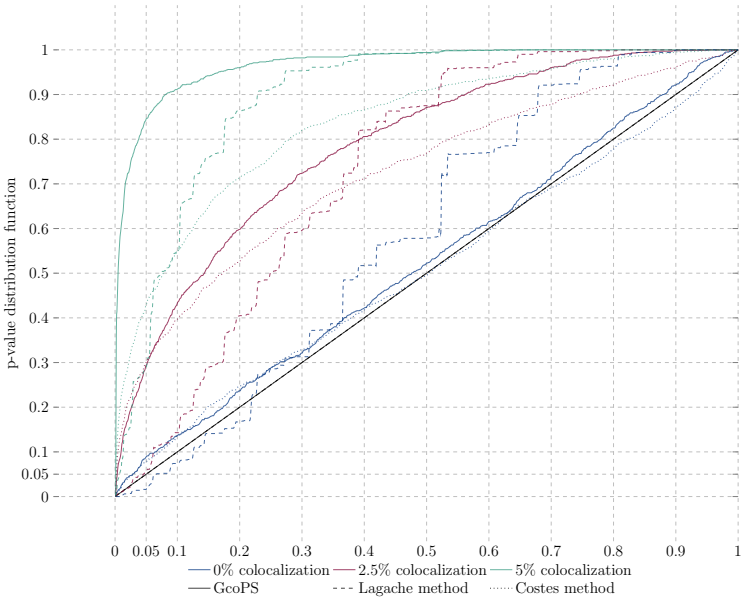
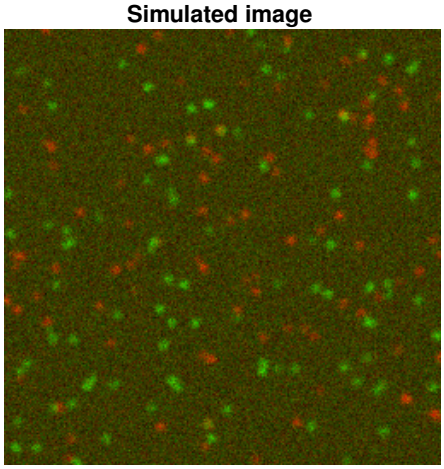


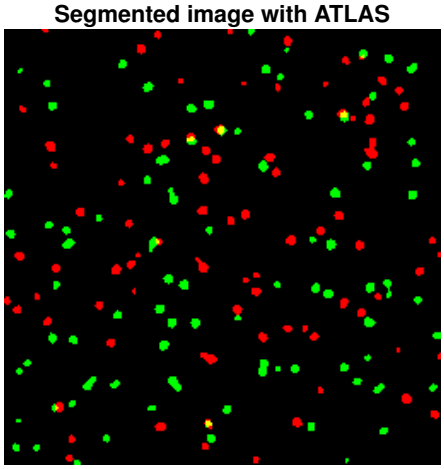
Fig. S1. Evaluation of GcoPS on 2D simulated images.

S2-a:

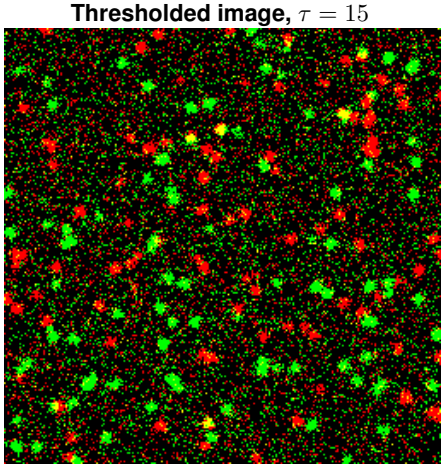
Simulated image



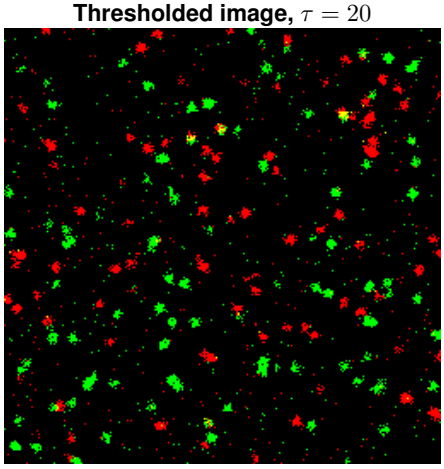
Segmented image with ATLAS



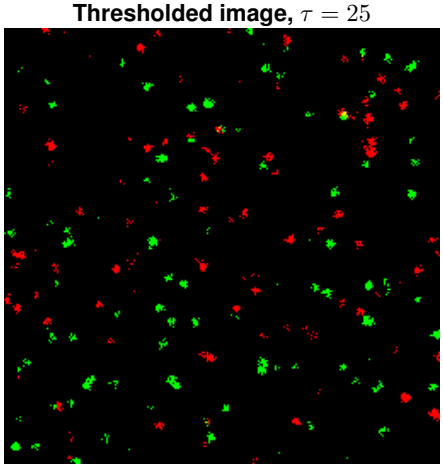
Thresholded image, $\tau = 15$



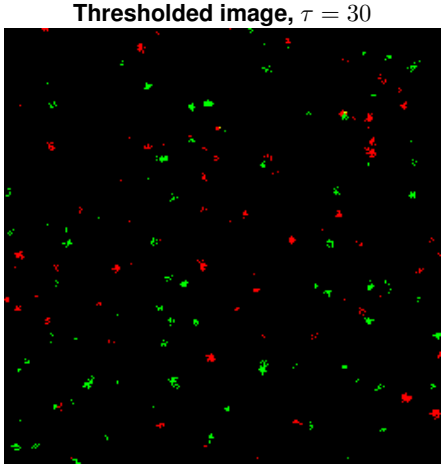
Thresholded image, $\tau = 20$



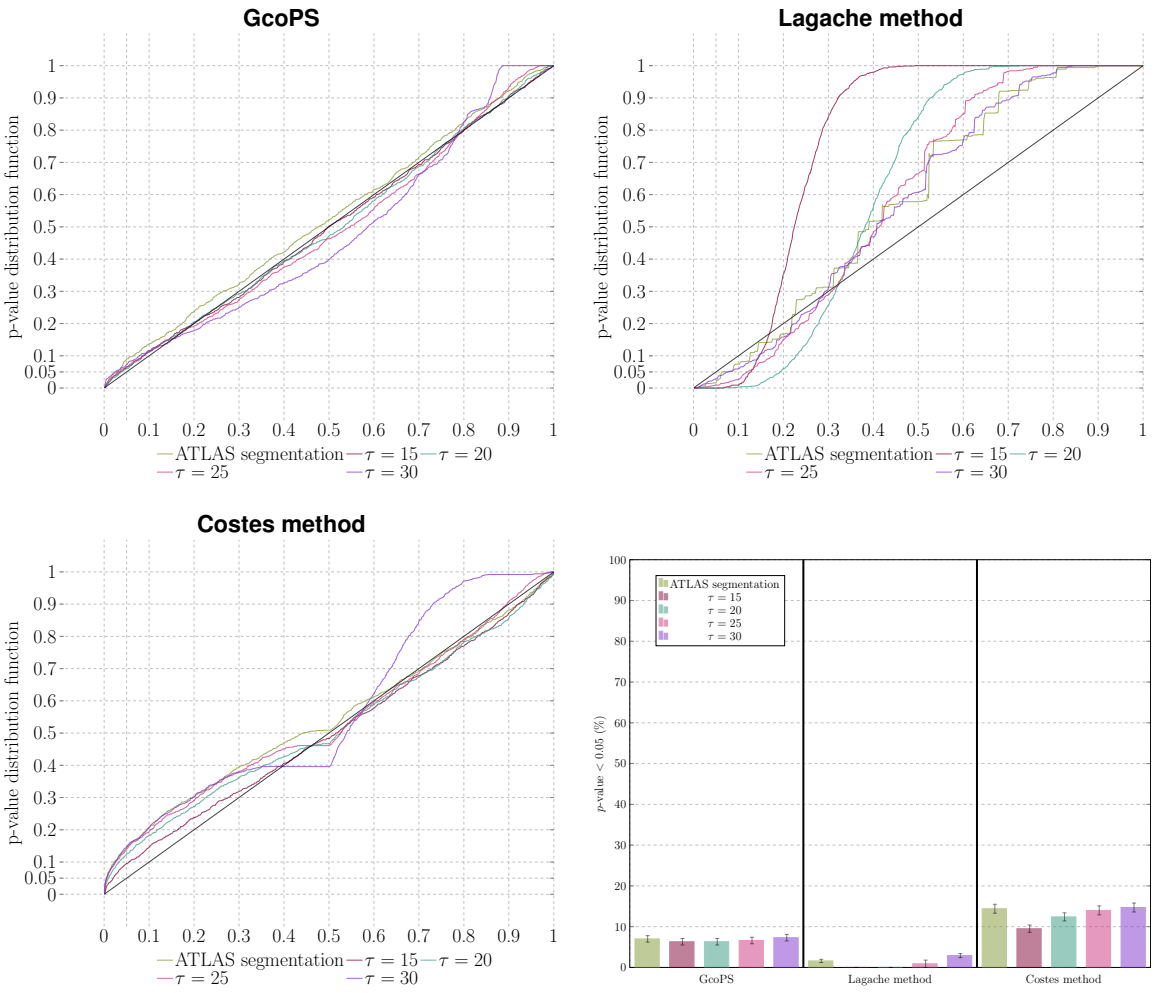
Thresholded image, $\tau = 25$



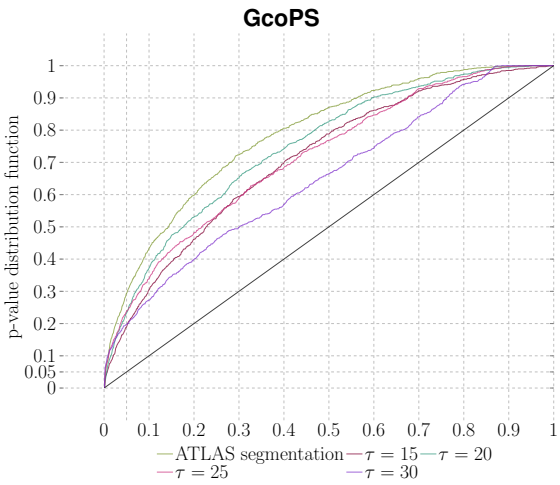
Thresholded image, $\tau = 30$

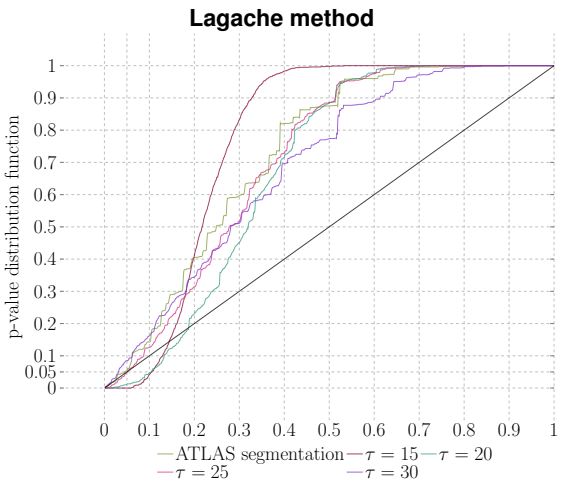


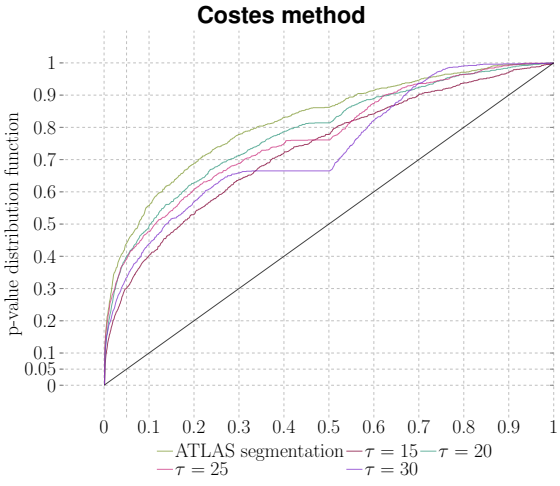
S2-b: 0% colocalization

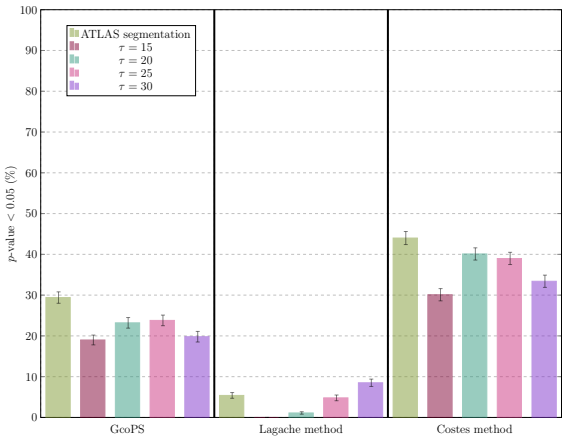


S2-c: 2.5% colocalization









S2-d: 5% colocalization

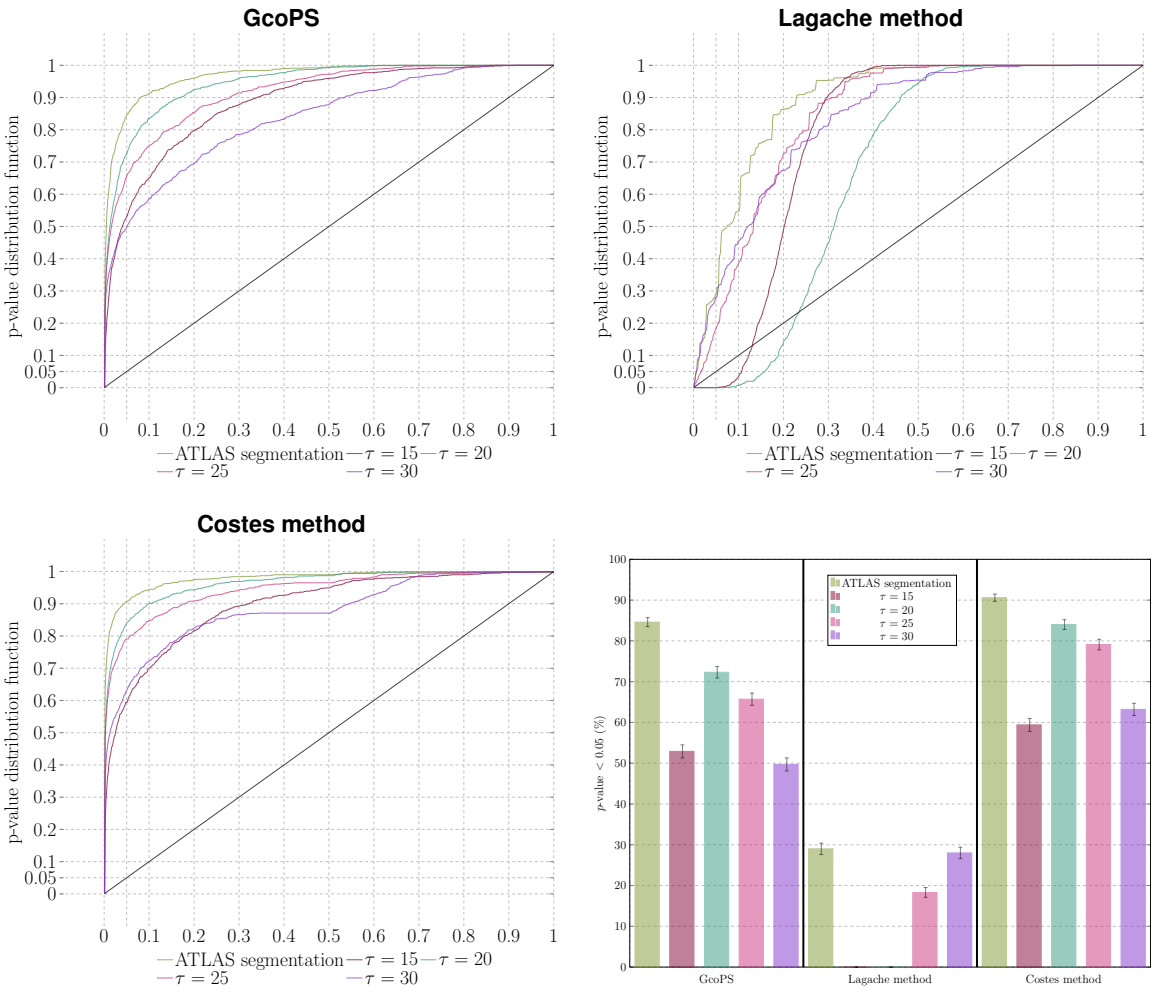
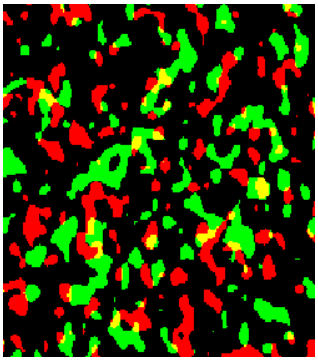
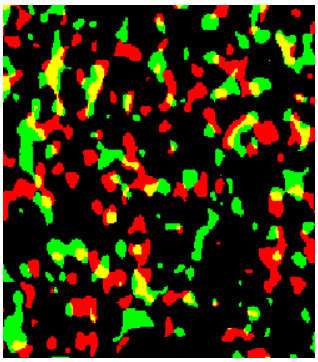


Fig. S2. Sensitivity to segmentation and deconvolution.

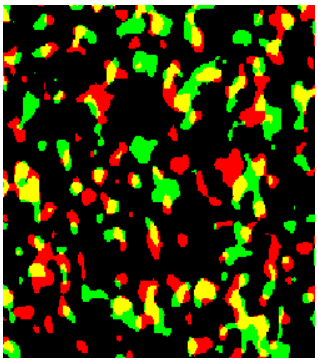
S3-a: Non regular objects



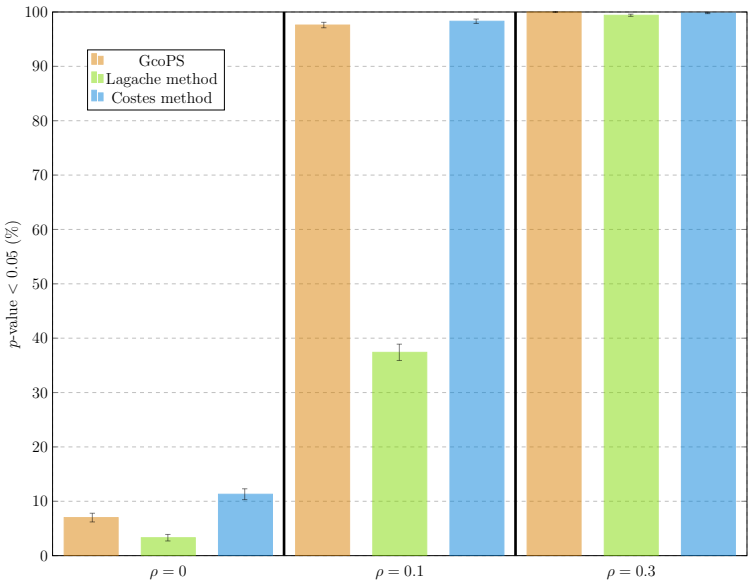
$\rho = 0$

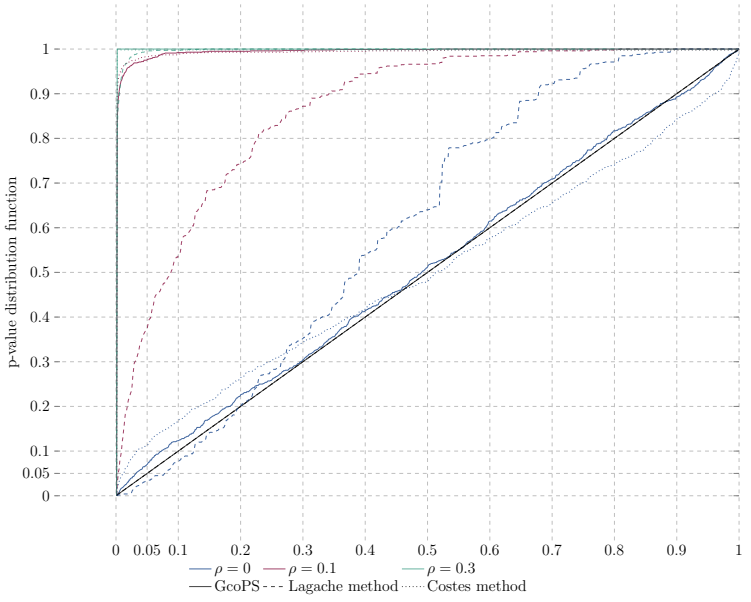


$\rho = 0.1$

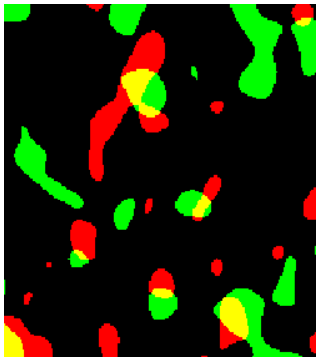


$\rho = 0.3$

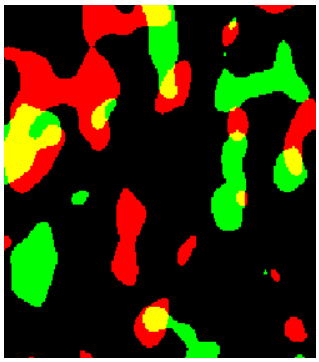




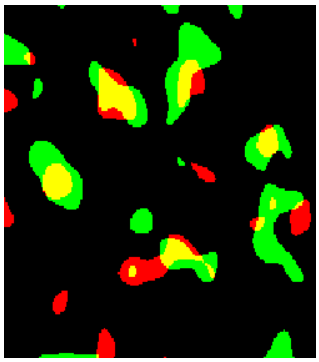
S3-b: Large and non regular objects



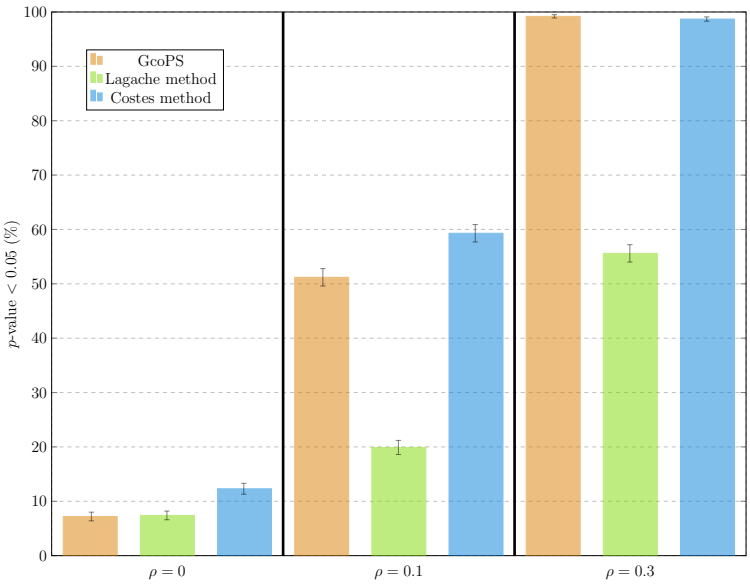
$\rho = 0$

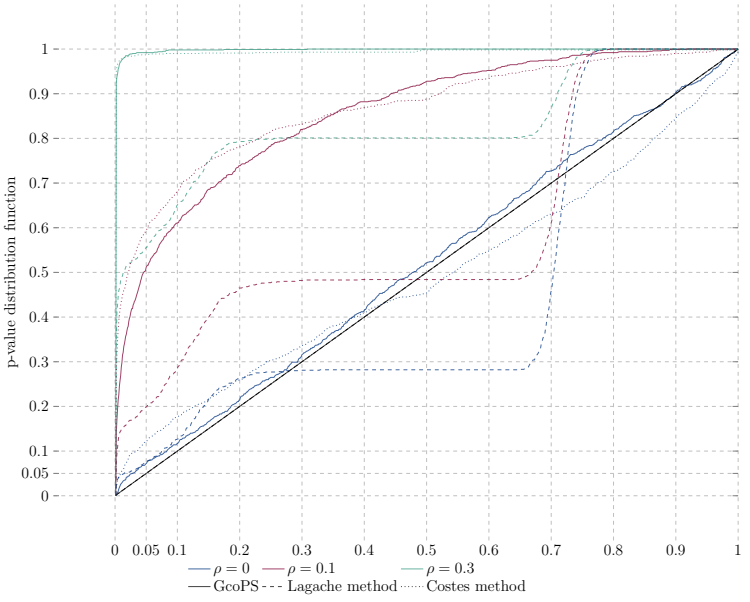


$\rho = 0.1$



$\rho = 0.3$



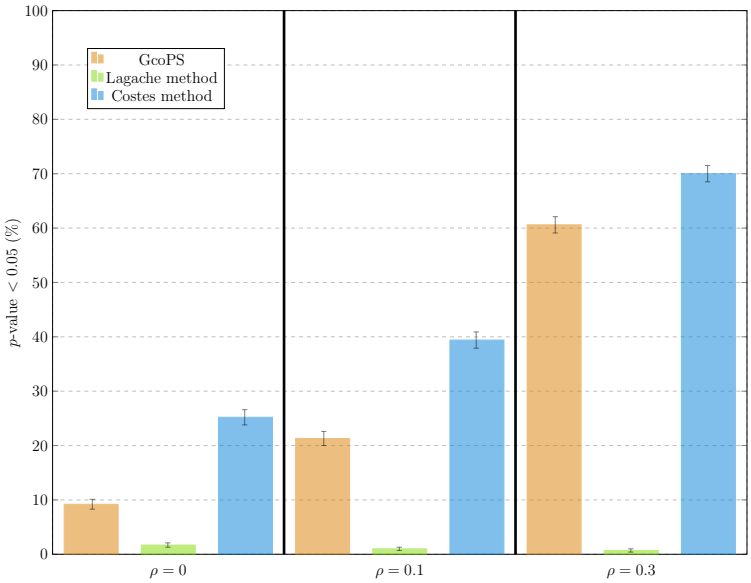


S3-c: Very large and non regular objects

$\rho = 0$

$\rho = 0.1$

$\rho = 0.3$



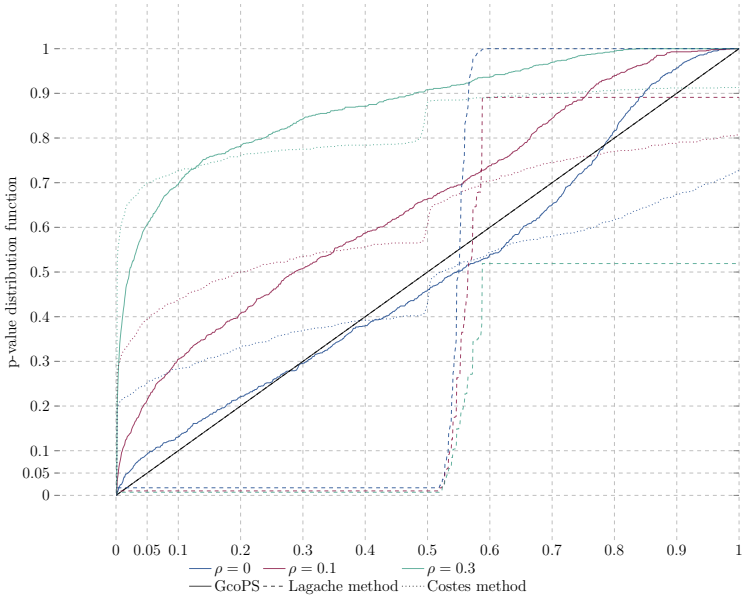
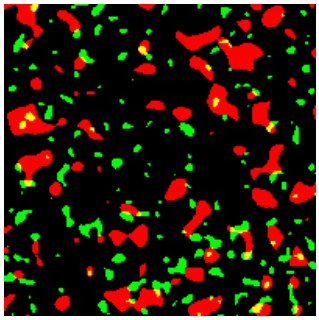
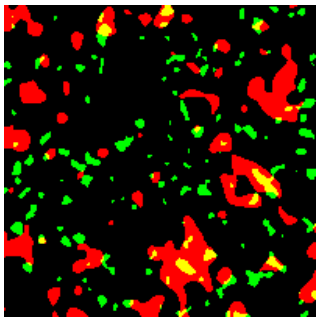


Fig. S3. Sensitivity to the shape and scale of objects.

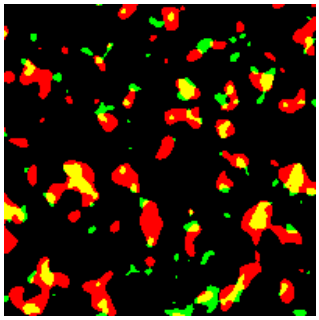
S4-a: Different resolutions in each channel



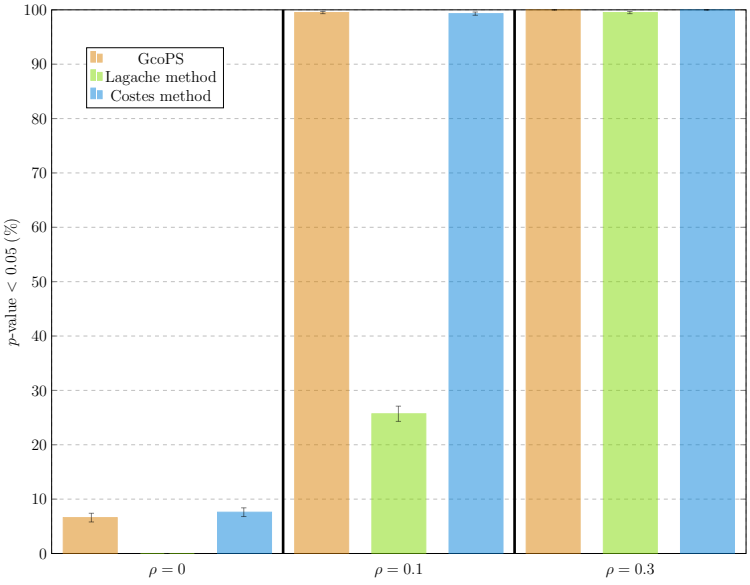
$\rho = 0$

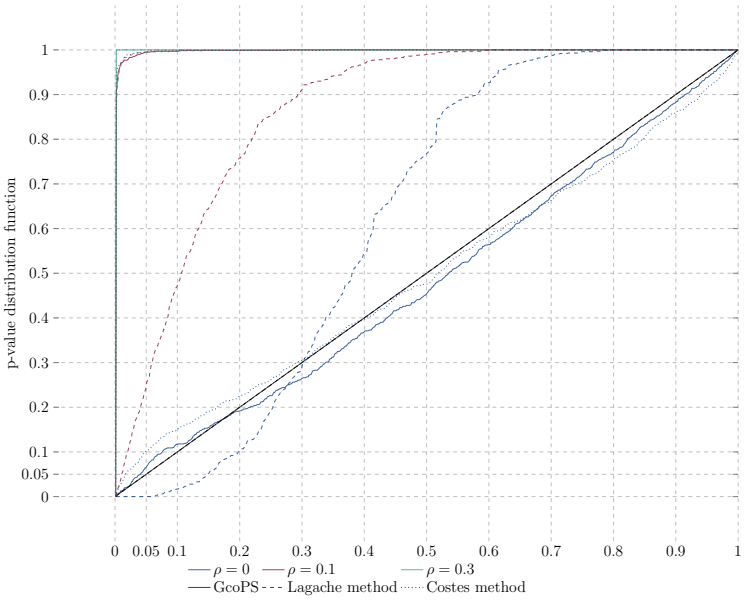


$\rho = 0.1$



$\rho = 0.3$





S4-b: Very different resolutions in each channel

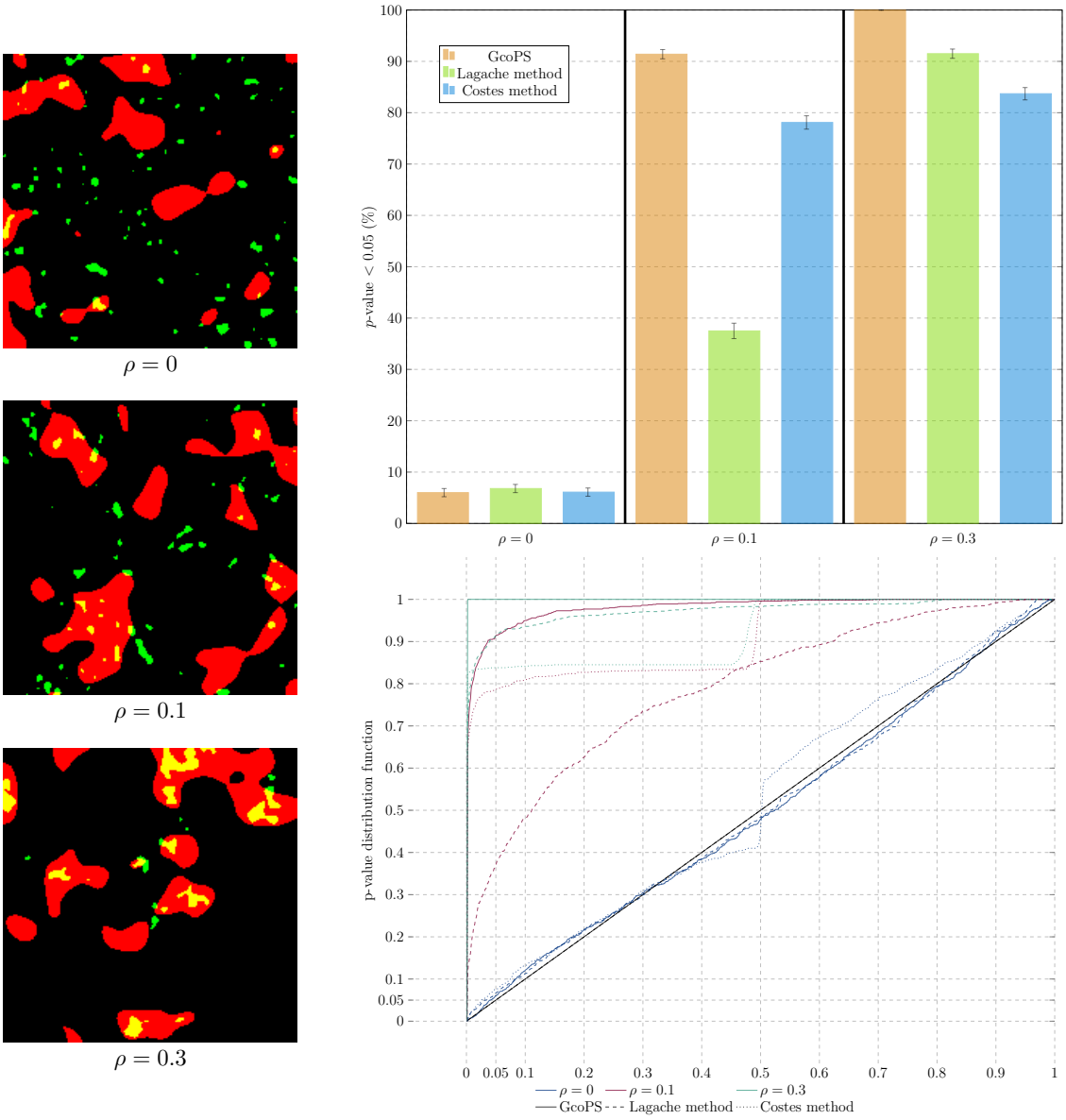


Fig. S4. Sensitivity to the resolutions in each channel.

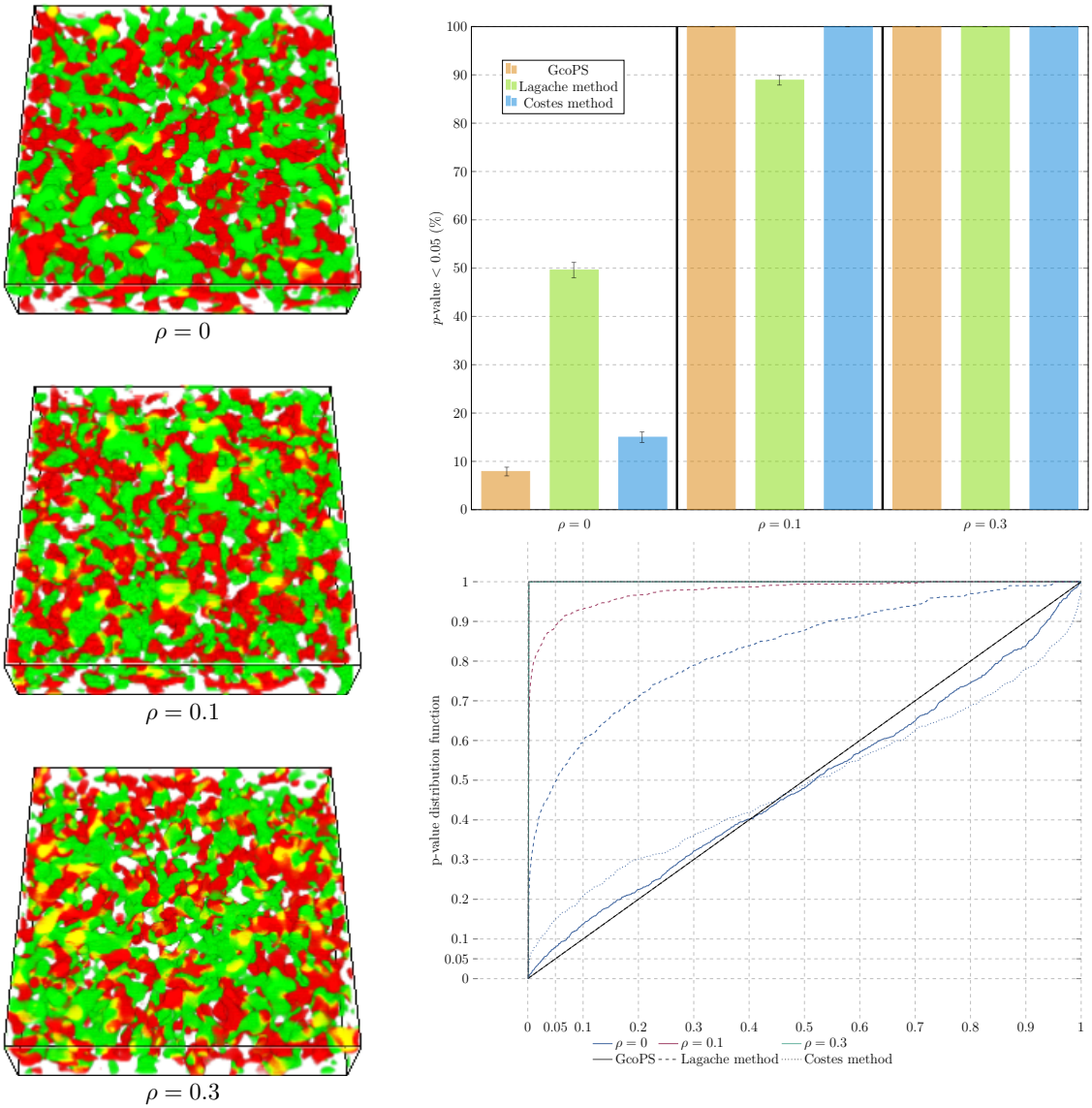


Fig. S5. Evaluation on 3D particles.

S6-a

S6-b

S6-c

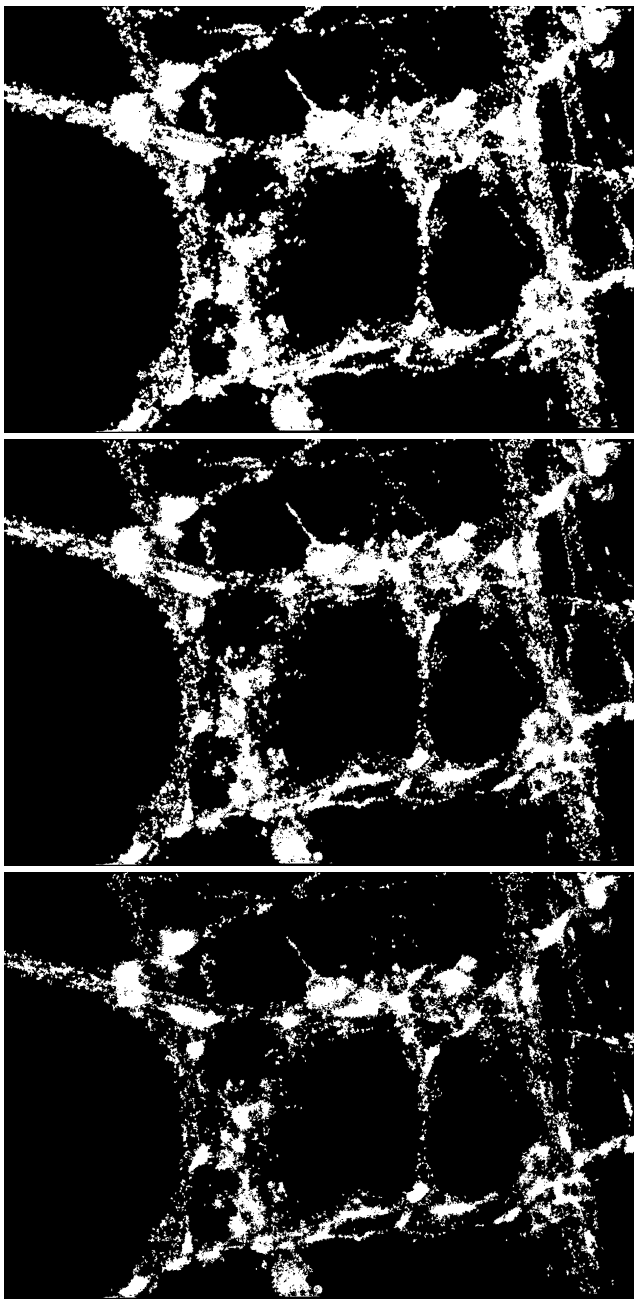


Fig. S6. vGlut segmentations.

	Costes[3]	**Lagache**[5]	**GcoPS**
CPU Time	★	★★★	★★★★
Sensitivity to method parameters	★	★★	★★★
Numbers of false positives	★	★★★★	★★★
Sensitivity to colocalization (true positives)	★★★★	★	★★★★
Robustness to segmentation outputs	★★★★	★★	★★★★
Robustness to non-regular shaped objects	★★	★	★★★★
Robustness to a different optical resolution in each channel	★★★	★★	★★★★

Table S1. Comparative summary.

	2D image 256×256 50 objects	**2D image** 256×256 200 objects	**2D image** 256×256 3500 objects	**2D+time image** $256 \times 256 \times 1000$ 100 objects	**3D image** $256 \times 256 \times 60$ 1000 objects	**3D image** $256 \times 256 \times 60$ 2000 objects
Costes[3] ImageJ plugin	6.1 sec	6.2 sec	6.1 sec	38 min 20 sec	3 min 3 sec	3 min 10 sec
Lagache[5] Icy plugin	1 sec	1.96 sec	12.38 sec	12 min 39 sec	25 sec	60 sec
GcoPS C++ code	**0.18 sec**	**0.2 sec**	**0.19 sec**	**29.5 sec**	**10 sec**	**9.8 sec**
GcoPS Icy plugin	**0.77 s**	**0.86 sec**	**0.82 sec**	**2 min 50 sec**	**22 sec**	**21 sec**

Table S2. Evaluation of computation time.