THE JOURNAL OF PHYSICAL CHEMISTRY B
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCB

Article

Critical Assessment of Condensate Boundaries in Dual-Color Single Particle Tracking

Published as part of The Journal of Physical Chemistry virtual special issue "Xiaoliang Sunney Xie Festschrift".

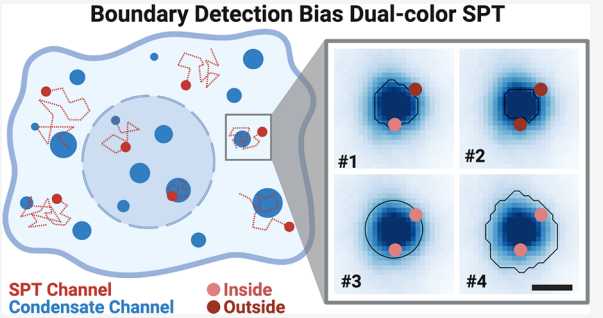
Guoming Gao and Nils G. Walter*

Cite This: *J. Phys. Chem. B* 2023, 127, 7694–7707

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: Biomolecular condensates are membraneless cellular compartments generated by phase separation that regulate a broad variety of cellular functions by enriching some biomolecules while excluding others. Live-cell single particle tracking of individual fluorophore-labeled condensate components has provided insights into a condensate's mesoscopic organization and biological functions, such as revealing the recruitment, translation, and decay of RNAs within ribonucleoprotein (RNP) granules. Specifically, during dual-color tracking, one imaging channel provides a time series of individual biomolecule locations, while the other channel monitors the location of the condensate relative to these molecules. Therefore, an accurate assessment of a condensate's boundary is critical for combined live-cell single particle-condensate tracking. Despite its importance, a quantitative benchmarking and objective comparison of the various available boundary detection methods is missing due to the lack of an absolute ground truth for condensate images. Here, we use synthetic data of defined ground truth to generate noise-overlaid images of condensates with realistic phase separation parameters to benchmark the most commonly used methods for condensate boundary detection, including an emerging machine-learning method. We find that it is critical to carefully choose an optimal boundary detection method for a given dataset to obtain accurate measurements of single particle-condensate interactions. The criteria proposed in this study to guide the selection of an optimal boundary detection method can be broadly applied to imaging-based studies of condensates.



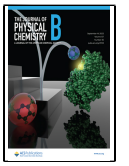
INTRODUCTION

Biomolecular condensates are cellular membraneless compartments ("organelles") formed by macromolecular phase separation accompanied by networking (gelation or percolation) or other physical transitions (henceforth referred to as PS ++),[1] which often leads to segregation of biochemical reactions in cells for differential regulation.[1,2] Starting from the discovery of their dynamic, liquid-like behavior in 2009,[3] biomolecular condensates have been found to play important roles in both physiology and pathology, mediating a broad range of cellular processes from gene expression to cell signaling.[1,4−7]

Live-cell single particle tracking (SPT) of protein or RNA molecules within condensates is a powerful approach to dissecting the organizational and functional impact principles of biomolecular condensates.[8−10] In particular, dual-color single particle tracking uses one channel to image micron-sized biomolecular condensates, usually via a fluorescently labeled marker protein, while using the other channel to follow the (typically subpixel-resolution) positions of single RNA or protein molecules over time to quantify their interactions with the condensate. The first channel thus needs to provide both the coordinates and the boundaries of each condensate to determine whether or not the biomolecule of interest resides within the condensate during a given movie frame. The time a single molecule spends within the condensate or dwell time, as well as the interaction kinetics between single molecules with the condensate, can be extracted from such cotracking. In addition, the series of locations of single molecules relative to the condensate centroid or boundary reveals their spatiotemporal organization and/or stoichiometry within the condensate.

The addition of a condensate channel to the single molecule channel is essential for interpreting the single molecule trajectories. Single-color SPT in live cells can provide information on the number of diffusion states of the labeled molecular species and the transition rates between these

Received: June 5, 2023
Revised: August 11, 2023
Published: September 5, 2023



https://doi.org/10.1021/acs.jpcb.3c03776
J. Phys. Chem. B 2023, 127, 7694−7707

states,[11,12] which may correlate with changing molecular interactions.[13] However, a change in diffusion state or the observation of anomalous diffusion alternatively can result from the formation of higher-order macromolecular complexes,[14] the emergence of constrained diffusion within a condensate environment,[15] the binding to other cellular components such as genomic DNA,[16−19] or the physical constraints imposed by surrounding intracellular membrane structures such as the Golgi apparatus or the endoplasmic reticulum.[20,21] Without direct evidence of a single molecule colocalizing with a condensate through a second imaging channel, it is difficult to draw clear conclusions.

Dual-color SPT of RNAs in RNP condensates or granules, termed processing bodies (PBs), has revealed both immobilized (static) and diffusing RNA molecules,[8] while similar observations were made when performing SPT of RNA-binding proteins (RBPs) in stress granules (SGs),[22] suggesting a partitioning of RNA and/or protein molecules and thus the existence of multiple phases within condensates. The SPT of RNAs in PB condensates additionally revealed distinct transient and stable interaction modes, dependent on the functional state of a given RNA.[8] The SPT of mRNAs in SGs further revealed that translation can occur within the condensates,[10] contradicting the long-standing assumption that SGs are incompatible with translation. Three-color SPT further revealed the interaction kinetics of RNAs with both PBs and SGs.[9,23] SPT approaches therefore should be able to address the question of whether RNP granules have biological functions or are simply "incidental condensates" emerging as nonfunctional byproducts of biomolecular concentration fluctuations in cells.[24]

An accurate measurement of single molecule dynamics relative to condensates relies on a correct assignment of the condensate's physical boundaries. Condensate boundary detection is especially challenging in dual-color SPT, because the number of fluorophores observed in the condensate channel is overwhelming compared to the SPT channel. The imaging laser power for the condensate channel therefore must be kept low to prevent bleed-through, reducing the signal-to-noise ratio in the condensate channel. More importantly, tools to quantitatively benchmark and compare the available boundary detection methods are missing. In this work, we first formulate the boundary detection challenge in dual-color SPT and rationalize the biases it introduces into measurements of single molecule-condensate interaction kinetics. We then generate simulated condensate images grounded in the physical model of phase separation to provide an absolute ground truth for benchmarking commonly used condensate boundary detection algorithms, similar to efforts applied to other single molecule biophysics tools.[25−28] A side-by-side comparison of four boundary detection methods based on our synthetic data across a wide physical parameter space yields important guidelines for choosing an optimal boundary detection algorithm for a given type of condensate. We suggest that these guidelines might also extend beyond dual-color SPT to provide guidance for other imaging-based analyses of biomolecular condensates.

METHODS

Live-Cell Dual-Color SPT. All examples of live-cell dual-color single molecule fluorescence imaging in this work are from SPT of a firefly luciferase (FL) messenger RNA (mRNA) labeled with Alexa Flour 647 (AF647) in the red channel, complemented with two types of Dcp1a-GFP condensates visualized in the green channel. The first type is PBs for which Dcp1a is a well-known marker,[8,29,30] and the second type is Dcp1a condensates formed via hyperosmotic phase separation (HOPS).[31,32]

For imaging both types of condensates, we used a previously characterized U2-OS cell line (hereafter referred to as UGD) that stably expresses GFP-tagged Dcp1a at near-physiological concentration so that a similar number and composition of Dcp1a-containing foci forms as in unlabeled parent U2-OS cells.[8] UGD cells were cultured at 37 °C in McCoy's 5A medium (Thermo Fisher, #16600082) with 10% (v/v) FBS (Fisher Scientific, #MT35016CV) and 20 U/mL Penicillin−Streptomycin (Invitrogen, #15140122). The cells were bead-loaded[33] with AF647-labeled FL mRNA, allowed to recover in culture medium for 1 h, and imaged in Leibovitz's L-15 medium without phenol red (Thermo Fisher, #21083027) but supplemented with 10% (v/v) FBS, all at 37 °C. Imaging in L-15 medium was preferred to balance any pH fluctuations caused by the cellular metabolism outside of a 5% CO_2 cell culture incubator, and the pH indicator phenol red was omitted to minimize fluorescence background. The imaging of PBs was performed under an isotonic condition, while Dcp1a HOPS condensates were induced under hypertonic conditions with the NaCl concentration raised from 150 to 300 mM. Both PBs and Dcp1a HOPS condensates were imaged on an Oxford Nanoimager at 5% power (∼0.1 mW) of the 488 nm laser, with an exposure time of 100 ms. The limited laser power prevents bleed-through of the green condensate channel into the red single molecule channel, while the 100 ms exposure time balances an adequate single-to-noise ratio against the diffusion motion blurring.

AF647-labeled FL mRNA was generated following a previous protocol.[34] Briefly, the FL encoding sequence is amplified by PCR from a pRL-CMV plasmid (Promega, Cat# E2261) while adding a T7 promoter sequence (5′-TAATACGACTCACTATAGGG-3′) in the upstream primer. The Kozak sequence and a 50-nt window upstream of the Kozak sequence were included to maximize translation initiation efficiency and leave space for a start-site recognition blind spot during the assembly of the human translation initiation complex.[35] RNA was transcribed from the PCR product using T7 RNA polymerase, purified by denaturing 7 M urea polyacrylamide gel electrophoresis, and subjected to a capping reaction (NEB, #M2080S), a polyadenylation reaction (Thermo Fisher, #74225Z25KU) first with 2′-azido-2′-dATP (Jena Biosciences, #NU-976S) then with rATP, and a click-chemistry labeling reaction with AF647 sDIBO alkyne (Thermo Fisher, #C20022) to produce translation-active, fluorophore-labeled FL mRNA. Labeling via click chemistry between the gene body and poly(A) tail was found to be the best strategy for generating functional in vitro transcribed mRNAs with minimal impact of fluorophore labeling on translation efficiency.[34,36] For live-cell dual-color SPT, the FL mRNA was imaged in the right channel of the Oxford Nanoimager, simultaneously with the condensates in the left channel at 8% power (∼2.0 mW) of the 640 nm laser and 100 ms exposure time. We used highly inclined optical light-sheet (HILO)[37] illumination to minimize out-of-focus fluorescence background.[38]

Simulation of Condensate Images. The simulation of condensate images follows a pipeline used previously in generate synthetic fluorescent microscopy images.[39] The

https://doi.org/10.1021/acs.jpcb.3c03776
J. Phys. Chem. B 2023, 127, 7694−7707

absolute ground truth for each simulated image is an idealized spherical condensate of radius R and with a partition coefficient (PC) given by the ratio of a condensate marker's concentration inside versus outside the condensate. Both R and PC are sampled by a low-discrepancy method called Sobol sampling, which has shown to be efficient in sampling a multidimensional parameter space.[40,41] R ranges from 100 to 600 nm, covering the size of PBs,[8,30] HOPS condensates,[31,32] and common RNP granules found in mammalian cells.[42] PC ranges from 2 to 10, matching the experimentally determined PC for common RNP granules.[43] Sobol sampling then ensures efficient sampling of the R−PC parameter space. The x and y coordinates of the condensate center are generated following a random distribution to place the condensate inside a 2 μm by 2 μm imaging field of view (FOV), with a padding of 600 nm to ensure that the FOV captures an entire condensate. Assigning a random location to the condensate minimizes bias from different pixel binning during down-sampling. Taken together, the absolute ground truth for each simulated condensate image can be expressed as:

$$F(x, y, R, \mathrm{PC}) \tag{1}$$

The ground truth image is generated from the absolute ground truth by factoring in other parameters arising from the common optical setups used for live-cell dual-color SPT. To simulate the image formation process, it is critical to perform a 3D convolution with an estimated point spread function (PSF). Although a mathematical estimation of a HILO illumination PSF is not available, it is well understood for a confocal microscope.[44] Live-cell dual-color SPT is commonly performed on both HILO[8,9,11,13,45,46] and spinning-disk confocal microscopes,[10,47−51] whereas confocal microscopes are more commonly used for bulk fluorescence detection requiring more homogeneous illumination across the FOV.[50] We decided to adopt the Gaussian approximation of a 3D PSF on the confocal microscope:[44]

$$G_{\sigma_{xy},\sigma_z}(x, y, z) = e^{-x^2+y^2/2\sigma_{xy}{}^2-z^2/2\sigma_z{}^2} \tag{2}$$

$$\sigma_{xy} = \frac{\sqrt{2}}{k_{\mathrm{em}}\mathrm{NA}} \tag{3}$$

$$\sigma_z = \frac{2\sqrt{6}\,n}{k_{\mathrm{em}}\mathrm{NA}^2} \tag{4}$$

$$k_{\mathrm{em}} = \frac{2\pi}{\lambda_{\mathrm{em}}} \tag{5}$$

where σ_{xy} is the standard deviation of the Gaussian on the x−y plane, σ_z is the standard deviation of the Gaussian on the z-axis, k_{em} is the emission wavenumber, NA is the numerical aperture, n is the refractive index, and λ_{em} is the emission wavelength. For this work, we assume that NA = 1.5, n = 1.515, and λ_{em} = 520 nm.

The condensate $F(x,y,R,\mathrm{PC})$ was placed into a simulation box with a voxel size of 10 nm, a width and length of 2 μm, and a height of $2(R + \sigma_z)$. The z-coordinate of the condensate was chosen as half the height of the simulation box. All voxels outside the condensate were assigned a fluorescence intensity of 1, with all voxels inside the condensate assigned a fluorescence intensity of 1 × PC. The whole simulation box was convolved with the Gaussian approximate of the 3D PSF on a confocal microscope, as rationalized above. After the 3D convolution, a z-slice with a thickness of 500 nm was taken out, representing the depth of focus (DOF) of the microscope. Any signal above or below the DOF slice of the PSF-convolved 3D box should not contribute to the image because it is out of focus. The DOF slice was projected onto the x−y plane by integrating all voxels above a pixel, resulting in a ground truth image with a pixel size of 10 nm. The ground truth image was then down-sampled to a noiseless synthetic image with a pixel size of 100 nm, mimicking fluorescent microscope images with a 100× objective. The down-sampling was done by averaging 10 by 10 pixels in the ground truth image. Finally, a mixture model was used to add noise to the noiseless synthetic image. The additive electronic noise was modeled by a Gaussian distribution, with parameters estimated from the experimental measurement of blank samples, and the shot noise from the integrating detector of the Oxford Nanoimager's sCMOS camera was modeled by a Poisson distribution.[52,53] A total of 4096 simulated images were generated. All python codes for this work, the four boundary detection methods, and the corresponding analyses and plots are found at https://github.com/walterlab-um/Condensate_Simulator_and_Boundary_Detector.

Boundary Detection by Algorithms Commonly Used for Condensate Images. We selected four commonly used boundary detection algorithms for benchmarking and comparison in this study. Please refer to the main text for a discussion of the published usage of each of the four methods.

Method #1 is thresholding after smoothing. The simulated image was smoothed by a σ = 1 Gaussian filter to reduce noise and then binarized by the Otsu thresholding algorithm.[54] Contours of the binary mask were found by the *findContours* function in the OpenCV package[55] and used as the detected boundaries for Method #1.

Method #2 is Canny edge detection,[56] which is a widely used edge detector in computer vision. The noise reduction step is the same as Method #1, but instead of applying simple thresholding to the smoothed image, the Canny algorithm performs hysteresis thresholding on the gradient of the smoothed image. In this work, Canny edge detection was done by calling the *Canny* function in OpenCV with the gradient calculated by convolving a size 5 Sobel filter and with low and high thresholds of 50 and 1000, respectively. The *findContours* function was then used to extract the detected boundaries for Method #2.

Method #3 is Laplacian of Gaussian (LoG) blob detection plus 2D Gaussian fitting. Method #3 is technically not a boundary detection algorithm but is included here because many biomolecular condensates in cells are small. The motivation is that objects of a few pixels do not have well-defined boundaries by traditional definition, which may cause edge detectors to fail, whereas a blob detector with a size estimator from Gaussian fitting may circumvent this obstacle. All the blobs in an image were first found using the *blob_log* function in the scikit-image package[57] with a maximum sigma value of 5. Briefly, the images were convolved by LoG kernels at different sigma values, where a blob will give the strongest response to LoG convolution when its size matches the size of the LoG kernel.[58] The center coordinates and sigma will then be used as initial guesses for 2D Gaussian fitting over a 7 by 7 window cropped around the center of the blob. The condensate boundaries of Method #3 were determined as 2.355 times the fitted sigma (i.e., the full width half maximum,

https://doi.org/10.1021/acs.jpcb.3c03776
J. Phys. Chem. B 2023, 127, 7694−7707

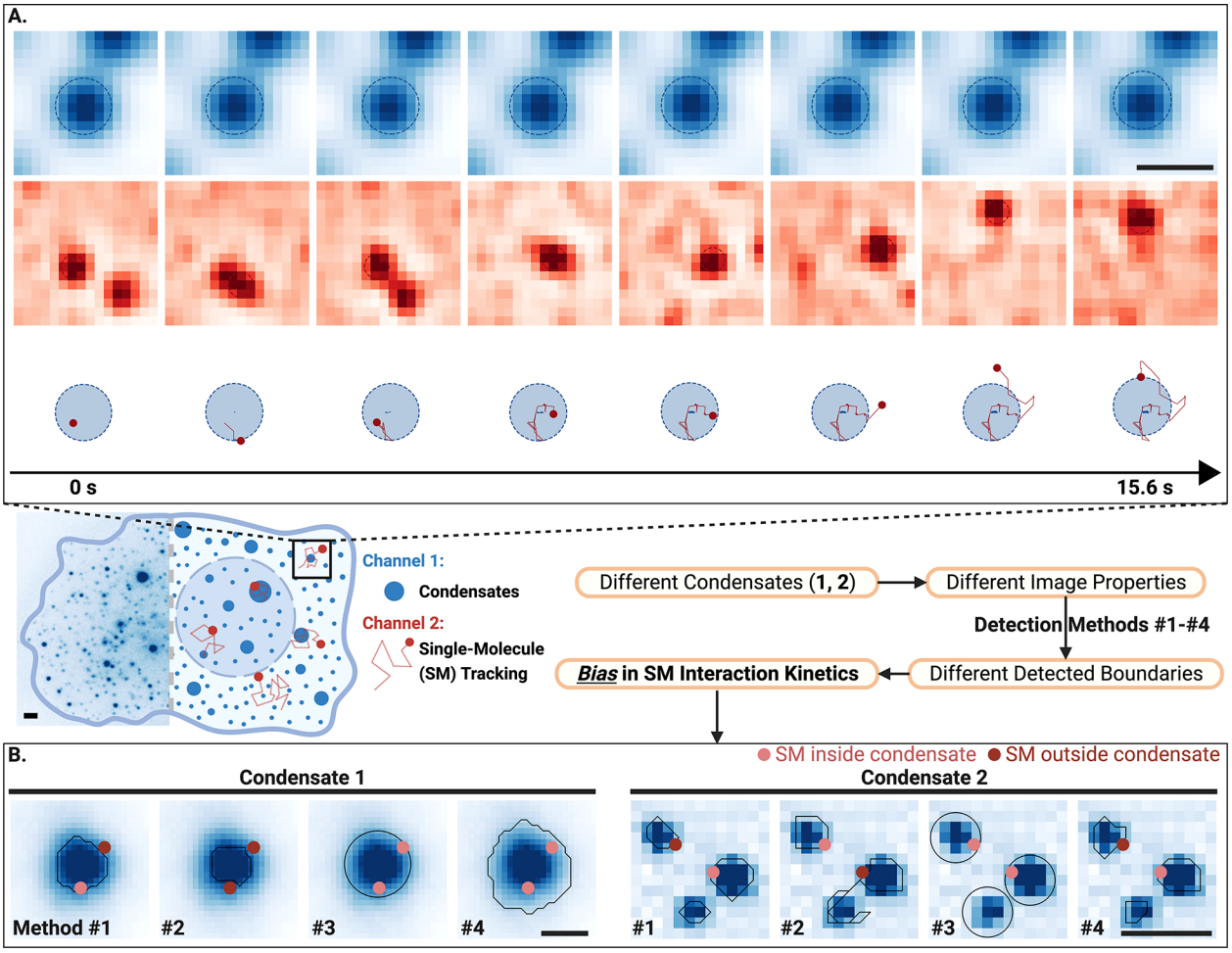


Figure 1. Accurate estimation of condensate boundaries is critical for dual-color single molecule tracking in biomolecular condensates. (A) Representative experimental single RNA molecule trajectory close to a HOPS condensate in a live UGD cell, with the condensates pseudo-colored in blue and the single molecule in red; the left half of the cell is an actual experimental image, the right half a schematic. The third row shows the reconstructed trajectory of the RNA relative to the condensate, with the estimated boundary of the condensate dotted in blue. (B) The performance of condensate boundary detection algorithms affects the single molecule-condensate interaction kinetics since it determines whether the single molecule is counted as inside or outside the condensate. Bias emerges broadly but differently depending on the type of condensate, as represented here by a large PB on the left and smaller Dcp1a HOPS condensates on the right. Methods #1−4 are the boundary detection algorithms benchmarked in this work, as detailed in the Methods section and in Figure 3. Scale bar, 1 μm.

FWHM, of a 2D Gaussian) away from the fitted center of the Gaussian.

Method #4 is a pixel classification by a user-trained machine-learning (ML) model. Condensates in an image were segmented by pixel classification, and the contours of the segmentation mask were used as the detected boundaries in Method #4. Although most publications using ML models for condensate images specify neither the training process nor the size of training dataset, a recent work using the ML model studied here, ilastik,[59] needed a training dataset of only seven images to achieve a reasonable recognition of the condensates. We generated a larger training dataset of 16 simulated condensate images and used Sobol sampling to ensure that the training dataset was representative of the R−PC parameter space. Seven independent researchers were first trained with a standard protocol for the well-established interactive ML tool ilastik[59] and then were given the above dataset to train the ML model to classify pixels into each a condensate and a background class. Of note, the researchers were not given the condensate ground truth and only used their naked eyes to make individual judgments of which pixel belongs to which class. This protocol was chosen to mimic the common practice of using ilastik on condensate images.[60] The trained model was then used to predict a segmentation mask under ilastik's "simple segmentation" mode in the export settings. Finally, the *findContours* function was called to extract the detected boundaries.

Metrics to Quantify the Goodness of Condensate Boundary Detection. The OpenCV package was used to extract centroid coordinates, area, and intensity within condensates, using in turn the boundary detected by each of the four methods. The detected PC was calculated as the average pixel intensity within the condensate boundary, divided by the average pixel intensity outside the condensate boundary. The deviation in condensate centers was derived as the physical distance between the contour centroid and the ground truth condensate center. The deviation in the condensate area and PC was calculated in fold change, as the difference between detected and ground truth values, divided by the ground truth value. The closer the fold change value is to one, the more accurate the measurement. The deviation in condensate boundary detection was computed as

https://doi.org/10.1021/acs.jpcb.3c03776
J. Phys. Chem. B 2023, 127, 7694−7707

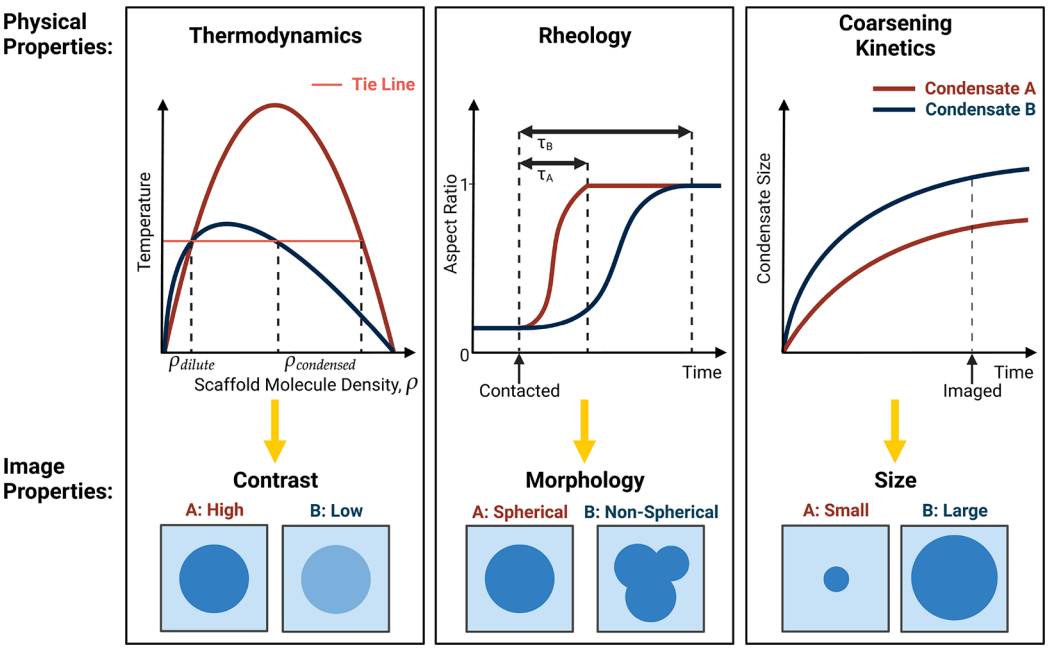


Figure 2. Physical origins of image properties of biomolecular condensates. The three major image properties of condensates, contrast, morphology, and size, can be traced back to three aspects of the physical processes that form biomolecular condensates in cells, namely, the thermodynamics, rheology, and coarsening kinetics of phase separation. Rheology and coarsening, relating to morphology and size in condensate images, are correlated because a faster coalescence means a faster Brownian motion coarsening. Depicted are the physical properties of two different condensates A and B together with their corresponding image properties. $\rho_{condensed}$ and ρ_{dilute} are the densities (i.e., concentrations) of the fluorescently tagged protein in the condensed and dilute phases, respectively. A tie line connects the densities of the two phases at equilibrium at a specific temperature. The aspect ratio is the ratio between the orthogonal short and long axes of a deformed condensate. τ_A and τ_B are the rheological relaxation times, defined as the time it takes for two fusing condensates to relax back to one condensate of spherical shape, that is, the recovery of an aspect ratio of ~1.

the root mean squared displacement (RMSD) between the detected and ground truth condensate boundaries. Each detected boundary coordinate was first used to calculate its distance to the ground truth condensate center. The distance minus the ground truth radius of the condensate was defined as a displacement to calculate the RMSD.

RMSD was chosen as the evaluation metric for condensate boundary detection rather than those used in evaluating segmentation algorithms,[61] for two major reasons: (1) All analyses for a dual-color SPT dataset, such as dwell kinetics and diffusion dynamics, are dependent on the determination of physical distances between single particle and condensate. The evaluation metric is therefore best chosen as a physical distance rather than a score without physical meaning. (2) Segmentation score functions are generally based on the comparison of a binary segmentation mask against a binary ground truth mask and thus have intrinsic pixel-level resolution. In contrast, the single molecule channel in SPT has subpixel resolution, as does the ground truth generated here by simulation, making RMSD a more straightforward and useful metric for evaluating SPT datasets.

RESULTS AND DISCUSSION

Boundary Detection Methods Generally Bias the Measurement of Single Molecule-Condensate Interaction Kinetics. Live-cell dual-color SPT aims to quantify the interactions between a protein or RNA molecule and a condensate over time. As illustrated in Figure 1A, the contour coordinates of the condensate boundary and the center coordinate of the single molecule will be extracted from each frame of the dual-color SPT video. The contour coordinates of the condensate boundary can be expressed in distinct ways, depending on the algorithm used. If assuming condensates are circles or ovals in the 2D image, the boundary coordinates are derived from the condensate's detected center coordinate and radius (Figure 1A). Otherwise, the edge detector or pixel classifier will directly yield boundary coordinates.

The resulting data can be used to reconstruct a single molecule trajectory relative to a condensate of interest, yielding rich information on the interactions between them. The third row of Figure 1A shows experimental data of a single RNA molecule that starts on the inside of a HOPS condensate in a live UGD cell, dwells in certain regions of the condensate, diffuses to another region to dwell there for a few seconds, repeats this jumping a few times, leaves the condensate briefly, and then comes back to enter the condensate. The regions of slow single molecule diffusion may represent subdomains that are more dense and viscous than other parts of the condensate, such as the core structures in SGs[22] or PBs.[8]

The relative locations of the RNA molecule and condensate throughout the trajectory require an accurate assessment of the condensate boundary. For the same condensate image from our live-cell dual-color tracking data, the four boundary detection algorithms most commonly used in the field yield different boundary contours so that the same single molecule in specific imaging frames may be counted as either inside or outside the condensate (Figure 1B). This observation holds broadly across different types of condensates, as represented here by larger PBs and smaller Dcp1a HOPS condensates (left and right segment, respectively, of Figure 1B). The bias across the four methods is partly dependent on the type of condensate, posing three significant questions for comparing

https://doi.org/10.1021/acs.jpcb.3c03776
J. Phys. Chem. B 2023, 127, 7694−7707

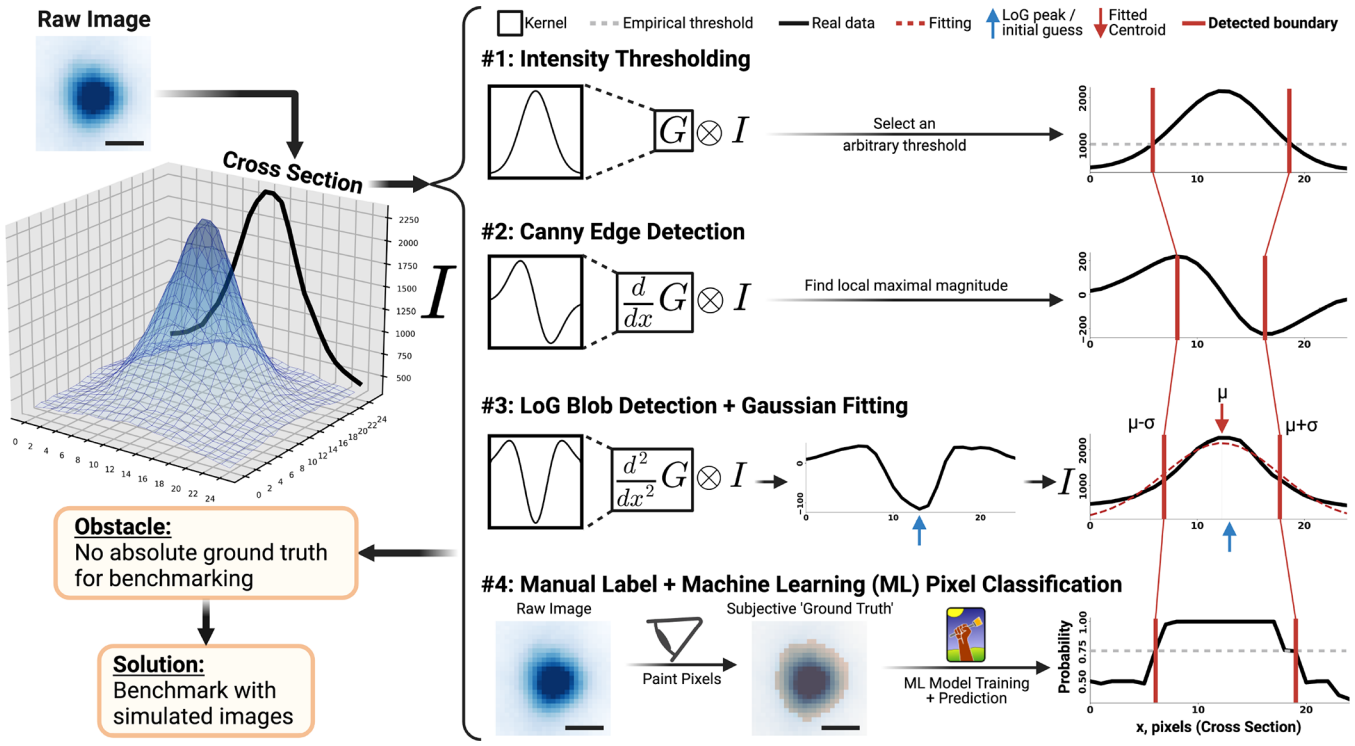


Figure 3. Common methods for condensate boundary detection and the obstacle for their objective comparison. The kernel for convolution is depicted on the left side of the convolution sign ⊗. For simplicity, the application of the four algorithms is shown on the one-dimensional (1D) cross-section of a condensate, whereas the analysis of real image data occurs in 2D. The Gaussian smoothing in Method #1 is sometimes substituted by convolution with other kernels. Scale bar, 1 μm.

them: (1) How large is the detection error (i.e., deviation or bias) in boundary detection? (2) How different are the detected boundaries? (3) How do the physical properties of condensates affect the detection error?

In the following sections, we first analyze the relationship between the physical properties and image properties of individual condensates. Second, we identify a lack of a ground truth as the major obstacle for an objective comparison between the four condensate boundary detection algorithms and propose synthetic data as a way to overcome this obstacle. Third, we discuss our findings and conclude with guidelines for choosing the best boundary detection algorithm based on the physical properties of a given type of condensate.

Image Properties of Condensates Originate from the Physical Parameters Governing Phase Separation. The formation of biomolecular condensates in cells is thought to be driven by phase separation, a density transition, possibly coupled with other changes such as in percolation behavior, a connectivity transition.[1,62] When taking an image or video of condensates, such as in dual-color fluorescence SPT experiments, the image properties of condensates can be traced back to their origin in the physical processes that induce condensates (Figure 2).

The image contrast of a condensate, defined as the ratio between the integrated pixel intensities within and outside the condensate, is determined by the concentration of fluorophores in each phase. Most dual-color single molecule-condensate tracking experiments use a condensate marker protein fused to a fluorescent protein to visualize the condensates so that the concentration of fluorophores equals that of the marker protein molecules. The concentration of the marker protein in each phase depends on its partition coefficient or PC. PS++ in cells is typically a multicomponent phase separation, although for simplicity we show a single-component phase diagram in the left column of Figure 2. Such a diagram is drawn with the x-axis representing an order parameter, density ρ or protein concentration, whereas the y-axis signifies a parameter affecting the enthalpic and entropic terms of the system, such as temperature. For a fixed temperature, if immiscibility occurs, the system at equilibrium separates into a condensed phase of density $\rho_{condensed}$ and a dilute phase of density ρ_{dilute}. In the phase diagram, these two equilibrium densities are connected by a tie line (Figure 2, left column). All the pairs of equilibrium densities adopted at different temperatures make up a coexistence or binodal curve. The system will then undergo phase separation if the combination of its total protein concentration and temperature lands within the region encircled by the binodal curve, with the tie connecting the two immiscible phases and the relative volumes of the two phases dictated by the distance of the system composition from the bimodal curve. In multi-component systems, the concentration of the marker protein in the condensed and dilute phases is derived from the phase diagram's x-axis. The PC is then simply the ratio of these concentrations, and the contrast of a condensate image increases with the PC of the fluorescently tagged protein marker, as derived from the phase diagram.

Once phase separation sets in, smaller condensates have a tendency to fuse (i.e., coalesce) to form entropically favored bigger condensates. The rheology of condensates will govern how fast two touching condensates can coalesce and relax back to the spherical shape that minimizes the new droplet's surface free energy.[63] The characteristic time for the latter process is termed relaxation time in rheology (τ_A and τ_B in Figure 2,

https://doi.org/10.1021/acs.jpcb.3c03776
J. Phys. Chem. B 2023, 127, 7694−7707

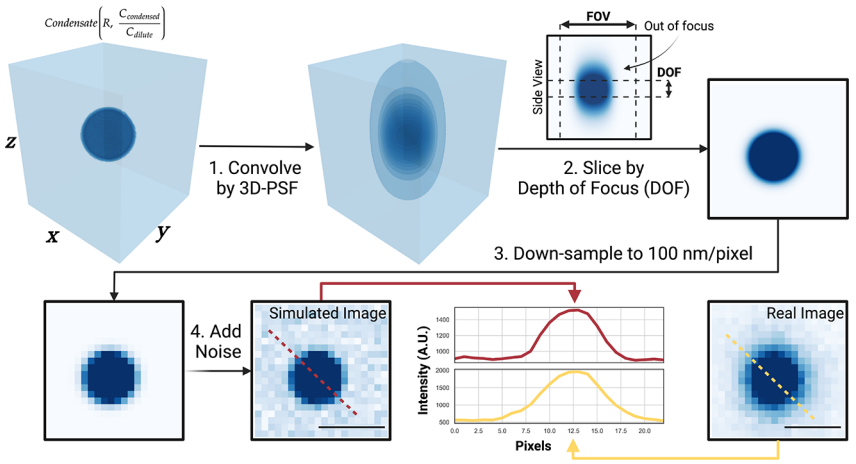


Figure 4. Simulated condensate images recapitulate real condensate images. The condensate is generated in a simulation box at the center z with only four parameters: the center coordinates x and y, size R, and the PC expressed as $C_{condensed}/C_{dilute}$, while assuming that the dilute phase concentration is fixed at a unitless 1. To simulate the image formation process, four steps are taken: (1) a 3D convolution with the PSF, (2) taking out a slice in z to mimic the DOF of the objective, (3) down-sampling to a representative pixel size, and (4) adding common imaging noise. Scale bar, 1 um.

middle column). If imaging takes place at a certain time point after two condensates start to fuse, the condensate with a shorter relaxation time (condensate A in Figure 2, middle column) will adopt the more spherical shape. Therefore, the morphology of condensates during imaging is derived from their rheology.

The increase in condensate size after nucleation and growth is termed coarsening, and it generally requires longer-range transport of or between condensates (Figure 2, right column).[64] Such long-range transport can be achieved by either Brownian motion of the condensates, termed Brownian motion coalescence (BMC), or diffusion of molecules from smaller to larger condensates, termed diffusion-limited coarsening (DLC) or Ostwald ripening.[64] The size of condensates at the time of imaging therefore depends on both the rheology, which determines the kinetics of coalescence, and the coarsening kinetics. Assuming that imaging takes place after the start of condensate assembly in a cell, condensates with faster coarsening kinetics (condensate B in Figure 2, right column) will be larger.

Taken together, the physical properties of biomolecular condensates of interest will determine their image properties. The contrast of condensates in the image will be determined by the PC, derived from the phase diagram and thus the thermodynamics. Condensate morphology or shape will be determined by the rheology. Finally, condensate size will be determined by the coarsening kinetics, which is affected by rheology if the coarsening mechanism is dominated by BMC. In the following sections, we will use condensate PC and size as the two major physical parameters to test how different condensates in an image are assessed by the various image processing algorithms.

Four Commonly Used Detection Algorithms for Condensate Boundaries. The literature on dual-color single molecule-condensate tracking and image-based analysis of biomolecular condensates yields four representative computer vision methods for the assessment of condensate boundaries (Figure 3). The following nomenclature of Methods #1−4 is used henceforth.

The most commonly used method is a direct thresholding of a smoothed image (Figure 3, Method #1).[9,10,65−68] Smoothing can be performed by convolution with a Gaussian kernel[10,23] as shown in Figure 3, a median kernel,[68] or a band-pass kernel.[9,65] Thresholding can be achieved either manually[65,66] or using an automatic thresholding algorithm such as the Otsu method[10,54] or the Yen method.[23,69] No matter how the threshold is picked, it should be considered arbitrary because it is derived without knowledge of the absolute ground truth.

A classical boundary detection in computer vision is the Canny edge detector (Figure 3, Method #2),[56] where the kernel used in convolution is the first derivative of a Gaussian instead of the Gaussian of Method #1. It has also been applied to condensate boundary detection.[70]

Method #3 assumes that condensates are blobs in the image so that the kernel used for convolution is the second derivative of a Gaussian (i.e., the Laplacian of Gaussian, LoG). The image is convoluted with an LoG kernel, which peaks at each blob that have the same size as the LoG kernel used.[71] Convolving the condensate image with a variety of LoG kernel sizes will then find the pixel locations of condensates of different sizes.[72] We note that peak picking relies on an arbitrary threshold set by users in LoG space.[66] 2D Gaussian fitting is then applied to find the subpixel locations and sizes of condensates,[9,65] as derived from the LoG blob detector (Figure 3, Method #3).

Lastly, pixel classifiers powered by machine-learning (ML) models such as ilastik[59] can be trained by a user-defined subjective "ground truth" to yield a mask where every pixel in the image is assigned to one of two categories, namely, condensate or background (Figure 3, Method #4).[60,73−79] The training of pixel classifiers can be convenient and quick. For example, ilastik is the most widely used interactive ML-based bioimage processing software has been used on condensate images, and it can work with as little as seven condensate images in the training dataset.[60]

All four methods rely on some empirical parameters and thresholds, and thus, the frame-to-frame changes in background texture or signal intensity can alter the overall performance of each method throughout a live-cell video. One solution this problem is using an automatic thresholding algorithm, such as the Otsu method[10,54] adopted in this paper, to find the best thresholds for each individual frame.

https://doi.org/10.1021/acs.jpcb.3c03776
J. Phys. Chem. B 2023, 127, 7694−7707

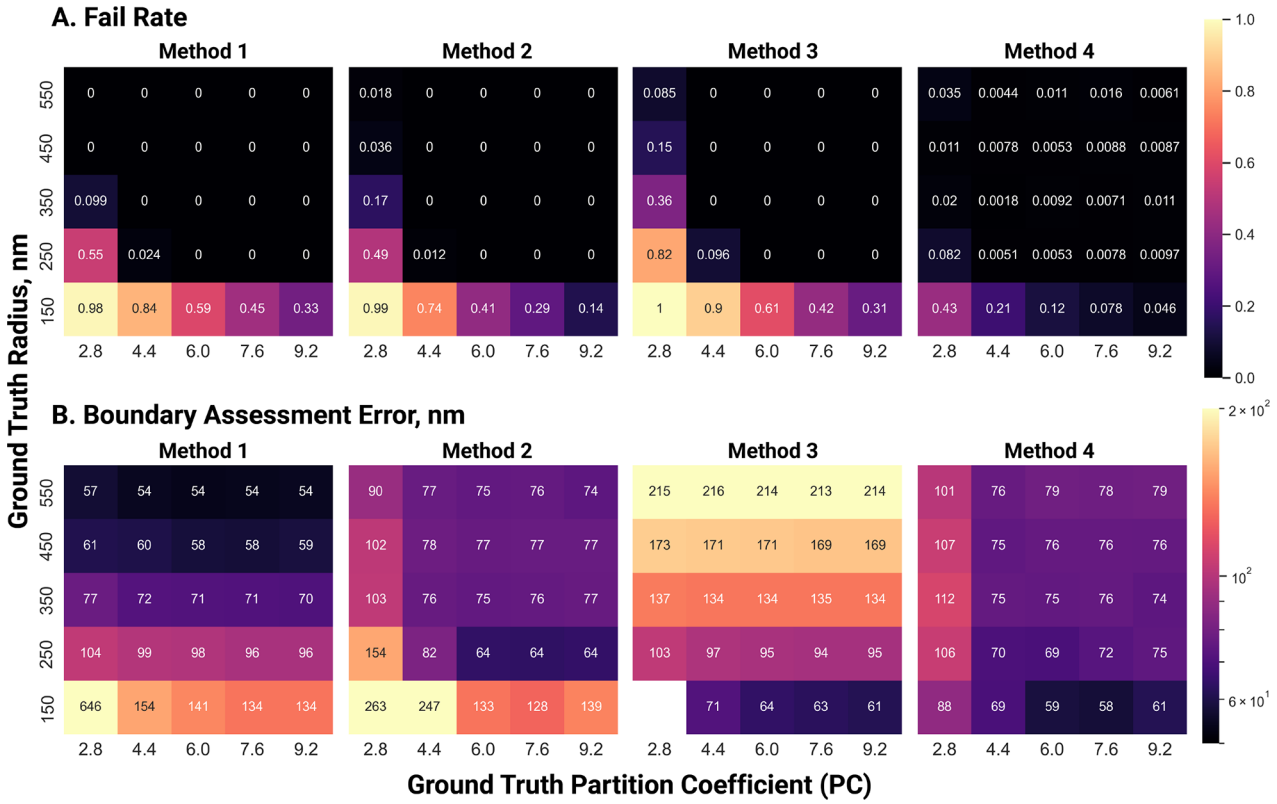


Figure 5. The best boundary detection method of condensates depends on both R (size) and PC (contrast). (A) Fail rate of condensate detection, defined as the likelihood of the method missing the condensate in an image. (B) Boundary deviation, defined as the RMSD distance between the detected edge and the ground truth edge. The darker the color, the smaller the fail rate and edge detection error, and thus the better a detection method performs in determining the condensate boundary.

Simulation Generates Synthetic Condensate Images that Define an Absolute Ground Truth. The biggest hurdle to benchmarking the performance of condensate boundary detection methods is the absence of a defined absolute ground truth. Here, we use simulations to generate synthetic condensate images based on two of the most important physical properties of condensates, their size and PC (Figure 4). These images lay the foundation for a quantitative comparison between the four boundary detection methods described above. Details of the simulations are found in the Methods section.

Our simulated images capture the major characteristics of a real condensate image (Figure 4). The remaining differences are either intrinsic to the complex cellular environment or due to the fact that the model has not yet considered the rheology of biological phase separation by assuming a spherical shape of the condensate. Condensate-prone proteins are known to form clusters with different stoichiometries at or below the saturation concentration[80,81] in the dilute phase outside the condensates, further increasing the heterogeneity of the image background. Additionally, the liquid-to-solid transition of condensates[82] and the presence of cellular structures like the cytoskeleton and membranes can deform a condensate to a nonspherical shape, even after the phase separation process has reached equilibrium. Future work could capture more complex condensate shapes by using either an empirical model, such as the parametric random shape-generating algorithm used in cell shape simulation,[83] or a simple physics model, such as the Ising model,[84] to simulate the phase separation process and to obtain condensate shapes without assuming equilibrium.

Optimal Condensate Boundary Detection Depends on Both Size and PC. The assessment of condensate boundaries starts with recognizing the condensate in an image. We first measured the failure rate of each method to pick up the condensate in a simulated condensate image. Condensates that are either small or of low PC were more likely to be missed by all four methods (Figure 5A). The classical computer vision Methods #1−3 nearly completely missed condensates that are smallest (radius 150 nm, or 3 pixels width in the simulated image) and have the lowest PC; however, the ML-based pixel classifier still recognized 55% of them (45% fail rate). Compared to the (nearly) perfect success rate (0% fail rate) of all four methods for larger and higher PC condensates (Figure 5A), the results call for caution when using any boundary detection methods on comparably small and dim condensates. An interesting observation is that the classical computer vision Methods #1−3 never miss a big and high-PC condensate, whereas the ML-based Method #4 has some finite fail rate for all types of condensates, including large- and high-contrast ones (Figure 5A). Notably, when a condensate is small but has a high PC (bottom row in Figure 5A) or has a low PC but is big (left-most column in Figure 5A), the ML-based Method #4 generally outperforms the classical computer vision Methods #1−3. Taken together, the ML-based Method #4 outperforms Methods #1−3 in recognizing most types of condensates, even if its fail rate is slightly higher for larger condensates of high contrast, which should be considered when an accurate counting of condensates matters.

To quantify the error in the condensate boundary detection, we calculated the boundary deviation as the RMSD distance

https://doi.org/10.1021/acs.jpcb.3c03776
J. Phys. Chem. B 2023, 127, 7694−7707

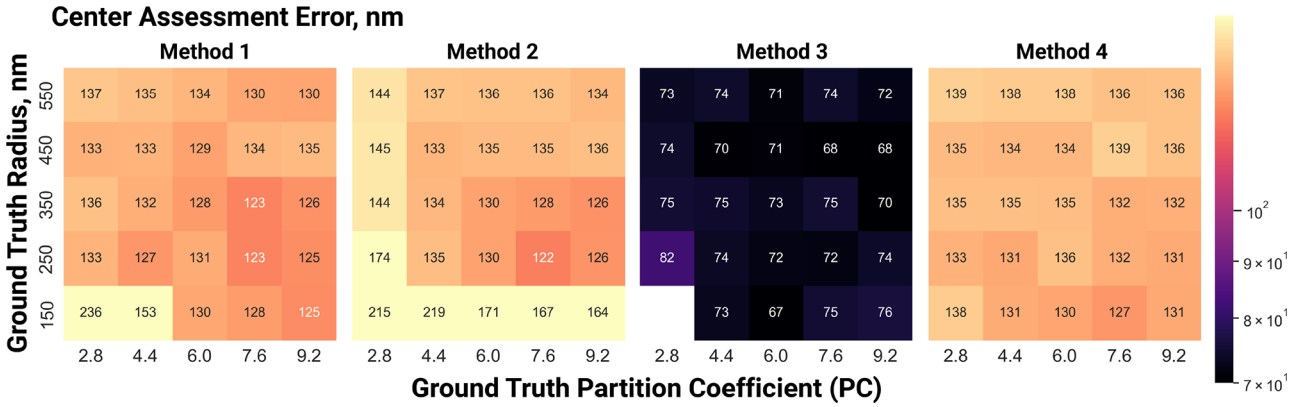


Figure 6. The LoG detector with Gaussian fitting best describes condensate center location. The condensate center deviation was defined as the distance between the detected condensate center and the ground truth center. The darker the color, the smaller the center detection error, and thus the better a detection method performs in determining the condensate location.

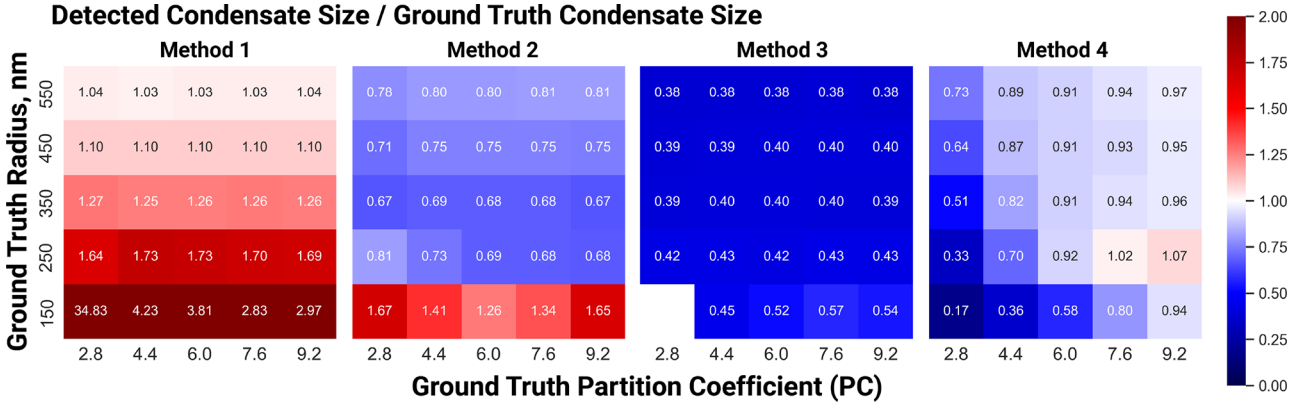


Figure 7. Different methods differentially bias the assessment of condensate size. The ratio between the detected condensate size over the ground truth condensate size quantifies how well a boundary detection method estimates the size of a condensate. The redder, the more an overestimation on condensate size occurs, while the bluer, the more an underestimation.

between the detected boundary and the known ground truth (Figure 5B). For large condensates (radius > ∼400 nm), the intensity thresholding Method #1 outperformed Methods #2−4. Therefore, Method #1 should be used if the condensates of interest in an experiment are large. In contrast, for very small condensates (radius < ∼200 nm), Methods #3 and #4 both displayed the highest boundary detection accuracies (Figure 5B). However, given the high fail rate of Method #3 (Figure 5A), Method #4 should preferably be employed if the condensates of interest are small. Finally, for medium-sized condensates (radius between ∼200 and ∼400 nm), the performance of Methods #1, #2, and #4 was similar, and when the PC of a condensate is less than ∼3.6, they all achieved an error of ∼100 nm (Figure 5B).

In summary, it is recommended to use Method #1 for condensates larger than ∼400 nm in radius and Method #4 for smaller condensates, or if the distribution of condensate sizes is broad, to achieve the most accurate condensate boundary detection in live-cell dual-color fluorescence SPT experiments that study the overall molecule-condensate interaction kinetics.

An LoG Detector with Gaussian Fitting Best Describes Small but Uniform Condensates. In some cases, the accurate assessment of a condensate's center may be more important than its boundary detection. For example, the estimation of a condensate's apparent diffusion coefficient relies on an accurate tracking of its center. When an SPT experiment is designed to measure single molecule interactions specifically with the cores of condensates, such as those of SGs,[85] the accuracy of the single molecule-to-condensate center distance will be more important than the single molecule-to-condensate boundary distance, and thus, the accuracy of condensate center assessment should be prioritized.

We compared the condensate center assessment errors, calculated as the deviation of the detected center from its ground truth, and found that Method #3 outperforms all other methods independent of condensate size and PC (Figure 6). This finding can be rationalized in that the LoG detector and Gaussian fitting are both designed to accurately locate blobs in an image. Therefore, Method #3 should be used for condensate center assessment in experiments where the single molecule-to-condensate center distance is important.

Different Methods Differentially Bias Estimation of Condensate Size. An accurate assessment of condensate size can be important for single molecule-condensate tracking because: (1) The distance between single molecule and condensate center or boundary may need to be normalized by the size of the condensate, especially if the condensates of interest have a broad size distribution. (2) It is sometimes necessary to normalize the single molecule-condensate colocalization probability to the total area or volume of condensates in the cell and relative to the total area or volume

https://doi.org/10.1021/acs.jpcb.3c03776
J. Phys. Chem. B 2023, 127, 7694−7707

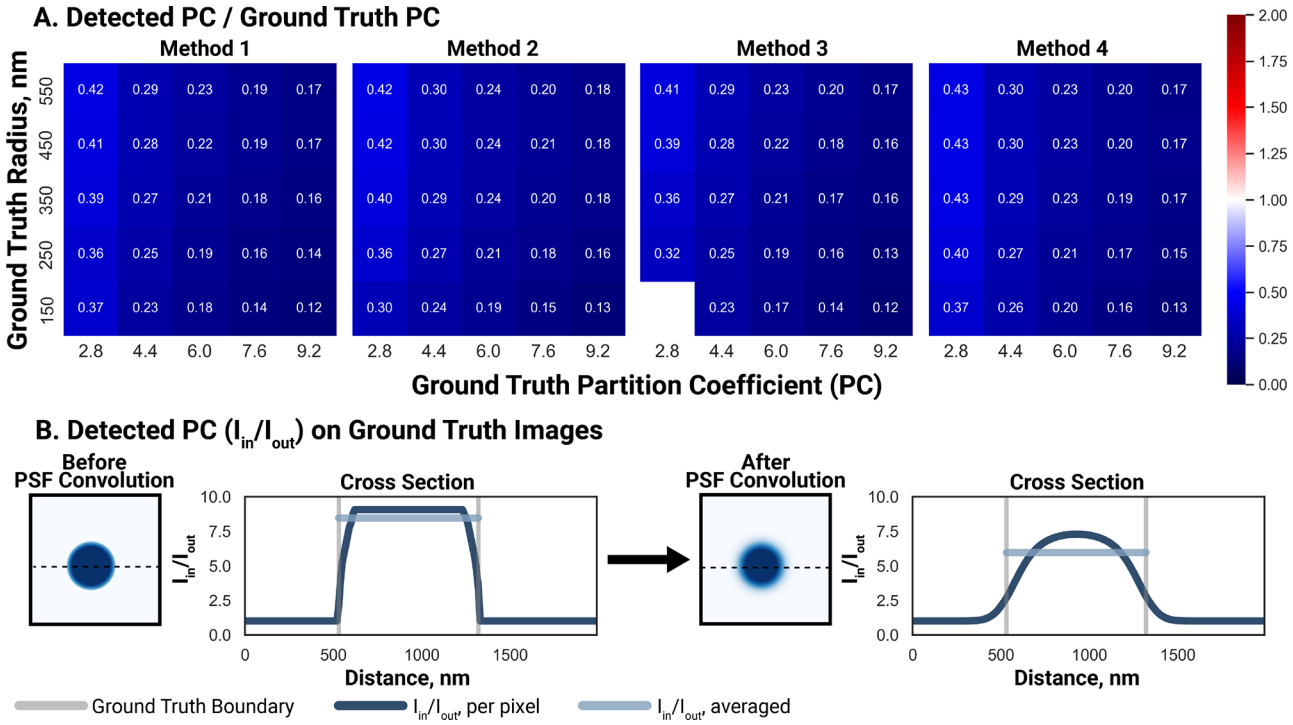


Figure 8. The PSF of fluorescence microscopy skews the image-based estimation of condensate PC. (A) All four methods severely underestimate PC. The ratio between the detected PC over the ground truth PC quantifies how well a boundary detection method estimates the PC of a condensate. The bluer, the more the PC is underestimated. (B) The bias of PC estimation originates from the PSF. Both images and cross-section curves are shown before down-sampling and background noise addition. The cross-section is taken from the dotted black line, cutting through the condensate in each image. The detected PC is expressed as the ratio between intensity within and outside the condensate. The ground truth boundary in gray is used to remove any impact of the boundary detection method used. The deep blue curve is the detected PC calculated by dividing intensity of each pixel by the average intensity outside the condensate, while the light blue curve is the detected PC averaged among all pixels inside the condensate, divided by the average intensity outside the condensate. The detected PC drops significantly after performing a convolution to mimic the PSF blurring intrinsic to fluorescence microscopy.

of the cell. The number of random single molecule-condensate colocalization events will be higher when the condensates occupy relatively more total area or volume of the cell, even if there are no specific interactions. (3) The size distribution of condensates itself can be a useful metric in quantifying the biological phase separation in cells.

Therefore, we quantify the ratio between the detected and ground truth condensate size using all four methods (Figure 7). Surprisingly, different methods will bias the assessment of condensate size in distinct ways. The intensity thresholding Method #1 yields a very faithful assessment of condensate size when the condensates are large (radius >∼400 nm) but will increasingly overestimate it with decreasing condensate size. Similarly, the Canny edge detection Method #2 underestimates condensate size when the ground truth size is large, while overestimating it when small. In contrast, the LoG detector plus Gaussian fitting Method #3 will generally underestimate condensate size as much as by 62%. This likely is caused by the fact that the shape of a large condensate in the ground truth image (Figure 4) is no longer close to the shape of a 2D Gaussian, which renders the fitting assumption increasingly inadequate when the ground truth condensate size increases. Finally, the overall performance of the ML-based Method #4 is better than Methods #1−3 but still falls short if the ground truth size and PC are small.

In summary, it would be best to use Method #1 for the size assessment of condensates with a radius larger than ∼400 nm. For smaller condensates, Method #2 or #4 is relatively better than the other two, but different biases can occur (Figure 7).

Imaging-Based Analysis Intrinsically Underestimates PC. We also quantify the error in the assessment of PC (Figure 8A). Unexpectedly, all four methods severely underestimate the PC regardless of the ground truth condensate size and PC, suggesting that the bias is intrinsically independent of the boundary detection process. One advantage of using synthetic data is that the entire image formation process can be broken down into individual steps, enabling us to understand the origin of the PC estimation bias. We found that the PSF of fluorescence microscopy skews the image-based estimation of condensate PC, independent of both the down-sampling to real pixel size and the background noise addition (Figure 8B). Therefore, any imaging-based estimation of PC using fluorescence microscopy image analysis will intrinsically underestimate the PC, which is significant for the field given that fluorescence microscopy is widely used to estimate the PC.

Annotator Bias of the Machine-Learning Model. Annotator bias refers to the tendency of human annotators to introduce subjective bias into training data, which has been a well-known and widely discussed problem in the ML field. However, it is rarely discussed when ML-based boundary detection tools like ilastik are used in biomolecular condensate research. Therefore, we compared the ML model trained by seven researchers given the same training dataset and instructions. We found that the distribution of the boundary

https://doi.org/10.1021/acs.jpcb.3c03776
J. Phys. Chem. B 2023, 127, 7694−7707

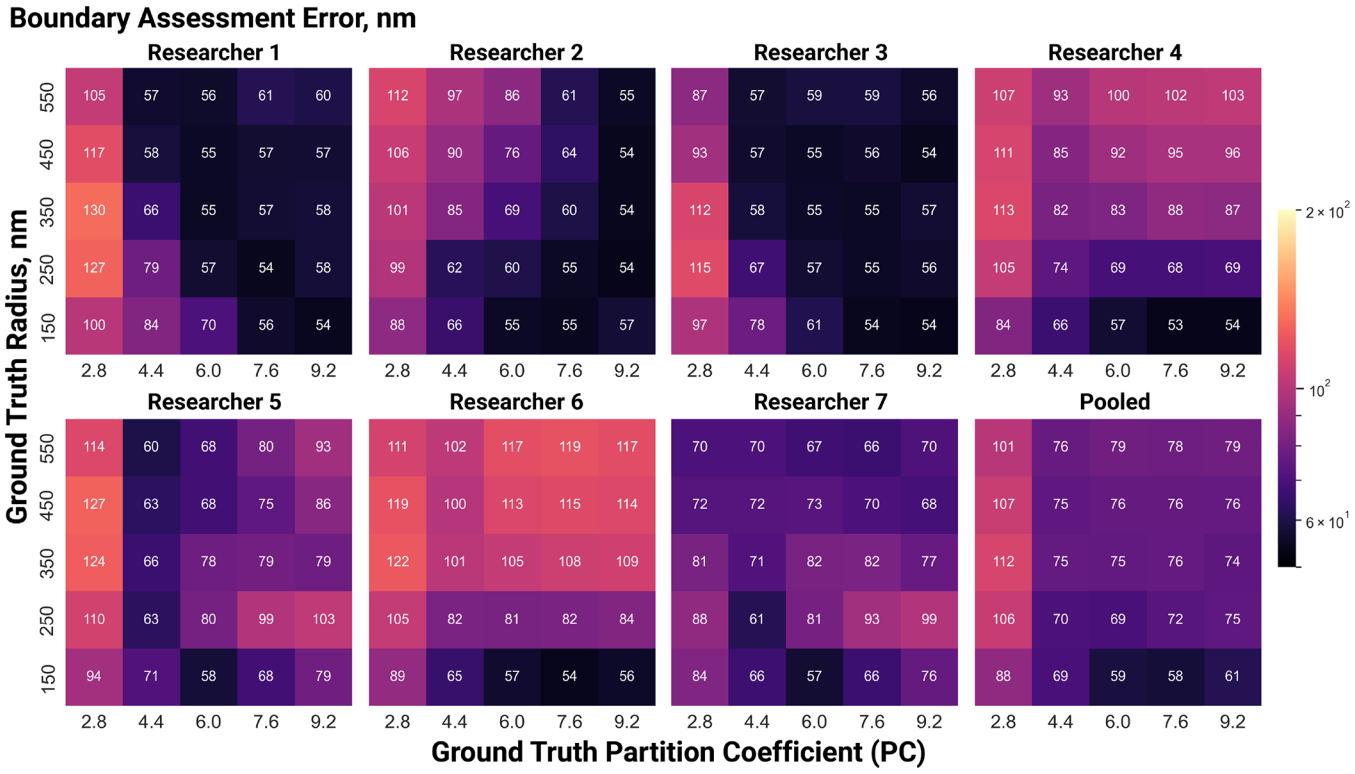


Figure 9. Annotator bias of ilastik on biomolecular condensates. The boundary detection error is defined as the RMSD distance of the detected boundary from the ground truth boundary. Models trained by different researchers perform differently in terms of condensate radius and PC. The darker the color, the better the ML model performs in detecting the condensate boundary defined by the ground truth.

detection is distinct among models trained by different researchers (Figure 9). Therefore, annotator bias also exists for the condensate boundary detection using ML-based methods and should be noted and properly addressed. For example, only experienced researchers should serve as annotators for condensate images, and experimental results could be cross-validated using ML models trained by different annotators.

CONCLUSIONS

Live-cell dual-color SPT of condensates is a powerful biophysical tool that (1) reveals the internal organization principles of condensates, with the single molecules serving as probes, and (2) elucidate the impact of condensates on the molecular functions of the RNA or protein. The single molecule-condensate tracking analysis relies on an accurate assessment of the condensate boundary.

Here, we analyzed the physical origins of image properties obtained during the image analysis of fluorescent condensates and used simulated condensate images as ground truth to subjectively benchmark four common condensate boundary detection methods: #1, intensity thresholding after smoothing; #2, Canny edge detection; #3, LoG blob detection plus 2D Gaussian fitting; #4, ML-based pixel classifier in ilastik. Based on the results, for the most accurate condensate boundary detection, we recommend to use Method #1 for condensates larger than ∼400 nm in radius and Method #4 for condensates smaller than ∼400 nm or when the condensate size has a broad distribution.

In some use cases, the assessment of the condensate center may be more important than the boundary. For example, some condensates may have a core−shell structure so that single molecule-condensate center distances may be more useful for the study of single molecule-condensate core interactions than the condensate boundary. Similarly, given that boundary detection is increasingly error prone for smaller condensates, it may be more accurate to use the single molecule-condensate center distance instead of the single molecule-condensate boundary distance for small condensates. In such cases, we suggest to Method #3 because it outperforms other methods in condensate center assessment across all sizes and PCs of condensates.

Both single molecule-condensate center distance and single molecule-condensate boundary distance can be normalized by the condensate size, since the condensate size distribution itself is a critical metric for biological phase separation. In this case, it would be best to use Method #1 for the size assessment of large condensates with a radius larger than ∼400 nm, whereas for smaller condensates, Methods #2 and #4 perform better than the other two, although different biases can occur depending on the choice of assessment method.

Finally, we note that any estimation of PC based on fluorescence microscopy images will underestimate the PC due to the physical nature of PSF blurring. In addition, the annotator bias inherent to ML models should be addressed properly when adopting Method #4.

AUTHOR INFORMATION

Corresponding Author

Nils G. Walter − *Center for RNA Biomedicine and Department of Chemistry, University of Michigan, Ann Arbor, Michigan 48109, United States;* orcid.org/0000-0002-7301-1275; Email: nwalter@umich.edu

https://doi.org/10.1021/acs.jpcb.3c03776
J. Phys. Chem. B 2023, 127, 7694−7707

Author

Guoming Gao − *Biophysics Graduate Program and Center for RNA Biomedicine, University of Michigan, Ann Arbor, Michigan 48109, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpcb.3c03776

Author Contributions

G.G. conceived the study, designed, performed all experiments and simulations, and analyzed the data. G.G. and N.G.W. wrote the paper. All authors have given approval to the final version of the manuscript.

Funding

This work was supported by NIH grant R35 GM131922 and a supplement on NIH grant R01 NS097542.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Liuhan Dai, Xiaofeng Dai, Emily Ellinger, Sarah Golts, Sujay Ray, Rosa Romero, and Emily Sumrall for help in training the ilastik pixel classifier. We also appreciate the helpful discussions of computer vision with Bangyan (Justin) Chu, as well as the fruitful discussions of phase separation with Xuemeng (Molly) Zheng. Lastly, we thank Ziyuan Chen, Laxmikanta Khamari, Chujie Sun, and Xuemeng (Molly) Zheng for their invaluable proofreading of and feedback on the manuscript. Figures 1−4 are created with BioRender.com.

■ ABBREVIATIONS

PS++	phase separation coupled to other transitions
PB	P body
RBP	RNA-binding protein
SG	stress granule
FL	firefly luciferase
mRNA	messenger RNA
HOPS	hyperosmotic phase separation
HILO	highly inclined laminated optical light sheet microscope
PC	partition coefficient
FOV	field of view
PSF	point spread function
DoF	depth of field
LoG	Laplacian of Gaussian
ML	machine-learning
RMSD	root mean squared deviation

■ REFERENCES

(1) Mittag, T.; Pappu, R. V. A conceptual framework for understanding phase separation and addressing open questions and challenges. *Mol. Cell* **2022**, *82*, 2201−2214.

(2) Pappu, R. V.; Cohen, S. R.; Dar, F.; Farag, M.; Kar, M. Phase Transitions of Associative Biomacromolecules. *Chem. Rev.* **2023**, *123*, 8945−8987.

(3) Brangwynne, C. P.; Eckmann, C. R.; Courson, D. S.; Rybarska, A.; Hoege, C.; Gharakhani, J.; Julicher, F.; Hyman, A. A. Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science* **2009**, *324*, 1729−1732. Research Support, Non-U.S. Gov't

(4) Alberti, S. Phase separation in biology. *Curr. Biol.* **2017**, *27*, R1097−R1102.

(5) Alberti, S.; Dormann, D. Liquid-Liquid Phase Separation in Disease. *Annu. Rev. Genet.* **2019**, *53*, 171−194.

(6) Lyon, A. S.; Peeples, W. B.; Rosen, M. K. A framework for understanding the functions of biomolecular condensates across scales. *Nat. Rev. Mol. Cell Biol.* **2021**, *22*, 215−235.

(7) Alberti, S.; Hyman, A. A. Biomolecular condensates at the nexus of cellular stress, protein aggregation disease and ageing. *Nat. Rev. Mol. Cell Biol.* **2021**, *22*, 196−213.

(8) Pitchiaya, S.; Mourao, M. D. A.; Jalihal, A. P.; Xiao, L.; Jiang, X.; Chinnaiyan, A. M.; Schnell, S.; Walter, N. G. Dynamic Recruitment of Single RNAs to Processing Bodies Depends on RNA Functionality. *Mol. Cell* **2019**, *74*, 521−533.e6.

(9) Moon, S. L.; Morisaki, T.; Khong, A.; Lyon, K.; Parker, R.; Stasevich, T. J. Multicolour single-molecule tracking of mRNA interactions with RNP granules. *Nat. Cell Biol.* **2019**, *21*, 162−168.

(10) Mateju, D.; Eichenberger, B.; Voigt, F.; Eglinger, J.; Roth, G.; Chao, J. A. Single-Molecule Imaging Reveals Translation of mRNAs Localized to Stress Granules. *Cell* **2020**, *183*, 1801−1812.e13.

(11) Hansen, A. S.; Woringer, M.; Grimm, J. B.; Lavis, L. D.; Tjian, R.; Darzacq, X. Robust model-based analysis of single-particle tracking experiments with Spot-On. *Elife* **2018**, *7*, No. e33125.

(12) Karslake, J. D.; Donarski, E. D.; Shelby, S. A.; Demey, L. M.; DiRita, V. J.; Veatch, S. L.; Biteen, J. S. SMAUG: Analyzing single-molecule tracks with nonparametric Bayesian statistics. *Methods* **2021**, *193*, 16−26.

(13) Hansen, A. S.; Amitai, A.; Cattoglio, C.; Tjian, R.; Darzacq, X. Guided nuclear exploration increases CTCF target search efficiency. *Nat. Chem. Biol.* **2020**, *16*, 257−266.

(14) Pitchiaya, S.; Androsavich, J. R.; Walter, N. G. Intracellular single molecule microscopy reveals two kinetically distinct pathways for microRNA assembly. *EMBO Rep.* **2012**, *13*, 709−715.

(15) Delarue, M.; Brittingham, G. P.; Pfeffer, S.; Surovtsev, I. V.; Pinglay, S.; Kennedy, K. J.; Schaffer, M.; Gutierrez, J. I.; Sang, D.; Poterewicz, G.; et al. mTORC1 Controls Phase Separation and the Biophysical Properties of the Cytoplasm by Tuning Crowding. *Cell* **2018**, *174*, 338−349.e20.

(16) Xie, X. S.; Choi, P. J.; Li, G. W.; Lee, N. K.; Lia, G. Single-molecule approach to molecular biology in living bacterial cells. *Annu. Rev. Biophys.* **2008**, *37*, 417−444.

(17) Biswas, S.; Chen, Z.; Karslake, J. D.; Farhat, A.; Ames, A.; Raiymbek, G.; Freddolino, P. L.; Biteen, J. S.; Ragunathan, K. HP1 oligomerization compensates for low-affinity H3K9me recognition and provides a tunable mechanism for heterochromatin-specific localization. *Sci. Adv.* **2022**, *8*, No. eabk0793.

(18) Fernandez, N. L.; Chen, Z.; Fuller, D. E. H.; van Gijtenbeek, L. A.; Nye, T. M.; Biteen, J. S.; Simmons, L. A. DNA Methylation and RNA-DNA Hybrids Regulate the Single-Molecule Localization of a DNA Methyltransferase on the Bacterial Nucleoid. *MBio* **2023**, *14*, No. e0318522.

(19) Chen, Z.; Seman, M.; Farhat, A.; Fyodorova, Y.; Biswas, S.; Levashkevich, A.; Freddolino, P. L.; Biteen, J. S.; Ragunathan, K. H3K9 methylation enhances HP1-associated epigenetic silencing complex assembly and suppresses off-chromatin binding. *bioRxiv* **2023**, DOI: 10.1101/2023.03.08.531771.

(20) Weiss, M.; Hashimoto, H.; Nilsson, T. Anomalous protein diffusion in living cells as seen by fluorescence correlation spectroscopy. *Biophys. J.* **2003**, *84*, 4043−4052.

(21) Wu, M. M.; Covington, E. D.; Lewis, R. S. Single-molecule analysis of diffusion and trapping of STIM1 and Orai1 at endoplasmic reticulum-plasma membrane junctions. *Mol. Biol. Cell* **2014**, *25*, 3672−3685.

(22) Niewidok, B.; Igaev, M.; Pereira da Graca, A.; Strassner, A.; Lenzen, C.; Richter, C. P.; Piehler, J.; Kurre, R.; Brandt, R. Single-molecule imaging reveals dynamic biphasic partition of RNA-binding proteins in stress granules. *J. Cell. Biol.* **2018**, *217*, 1303−1318.

(23) Wilbertz, J. H.; Voigt, F.; Horvathova, I.; Roth, G.; Zhan, Y.; Chao, J. A. Single-Molecule Imaging of mRNA Localization and Regulation during the Integrated Stress Response. *Mol. Cell* **2019**, *73*, 946−958.e7.

(24) Putnam, A.; Thomas, L.; Seydoux, G. RNA granules: functional compartments or incidental condensates? *Genes Dev.* **2023**, *37*, 354−376.
(25) Ulman, V.; Maška, M.; Magnusson, K. E. G.; Ronneberger, O.; Haubold, C.; Harder, N.; Matula, P.; Matula, P.; Svoboda, D.; Radojevic, M.; et al. An objective comparison of cell-tracking algorithms. *Nat. Methods* **2017**, *14*, 1141−1152.
(26) Muñoz-Gil, G.; Volpe, G.; Garcia-March, M. A.; Aghion, E.; Argun, A.; Hong, C. B.; Bland, T.; Bo, S.; Conejero, J. A.; Firbas, N.; et al. Objective comparison of methods to decode anomalous diffusion. *Nat. Commun.* **2021**, *12*, 6253.
(27) Chenouard, N.; Smal, I.; de Chaumont, F.; Maška, M.; Sbalzarini, I. F.; Gong, Y.; Cardinale, J.; Carthel, C.; Coraluppi, S.; Winter, M.; et al. Objective comparison of particle tracking methods. *Nat. Methods* **2014**, *11*, 281−289.
(28) Götz, M.; Barth, A.; Bohr, S. S.; Börner, R.; Chen, J.; Cordes, T.; Erie, D. A.; Gebhardt, C.; Hadzic, M.; Hamilton, G. L.; et al. A blind benchmark of analysis tools to infer kinetic rate constants from single-molecule FRET trajectories. *Nat. Commun.* **2022**, *13*, 5402.
(29) Aizer, A.; Brody, Y.; Ler, L. W.; Sonenberg, N.; Singer, R. H.; Shav-Tal, Y. The dynamics of mammalian P body transport, assembly, and disassembly in vivo. *Mol. Biol. Cell* **2008**, *19*, 4154−4166.
(30) Hubstenberger, A.; Courel, M.; Bénard, M.; Souquère, S.; Ernoult-Lange, M.; Chouaib, R.; Yi, Z.; Morlot, J. B.; Munier, A.; Fradet, M.; et al. P-Body Purification Reveals the Condensation of Repressed mRNA Regulons. *Mol. Cell* **2017**, *68*, 144−157.e5.
(31) Jalihal, A. P.; Pitchiaya, S.; Xiao, L.; Bawa, P.; Jiang, X.; Bedi, K.; Parolia, A.; Cieslik, M.; Ljungman, M.; Chinnaiyan, A. M.; et al. Multivalent Proteins Rapidly and Reversibly Phase-Separate upon Osmotic Cell Volume Change. *Mol. Cell* **2020**, *79*, 978−990.e5.
(32) Jalihal, A. P.; Schmidt, A.; Gao, G.; Little, S. R.; Pitchiaya, S.; Walter, N. G. Hyperosmotic phase separation: Condensates beyond inclusions, granules and organelles. *J. Biol. Chem.* **2021**, *296*, 100044.
(33) Cialek, C. A.; Galindo, G.; Koch, A. L.; Saxton, M. N.; Stasevich, T. J. Bead Loading Proteins and Nucleic Acids into Adherent Human Cells. *J. Visualized Exp.* **2021**, *172*, 1−16.
(34) Custer, T. C.; Walter, N. G. In vitro labeling strategies for in cellulo fluorescence microscopy of single ribonucleoprotein machines. *Protein Sci.* **2017**, *26*, 1363−1379.
(35) Querido, J. B.; Sokabe, M.; Kraatz, S.; Gordiyenko, Y.; Skehel, J. M.; Fraser, C. S.; Ramakrishnan, V. Structure of a human 48S translational initiation complex. *Science* **2020**, *369*, 1220−1227.
(36) Schmidt, A.; Gao, G.; Little, S. R.; Jalihal, A. P.; Walter, N. G. Following the messenger: Recent innovations in live cell single molecule fluorescence imaging. *Wiley Interdiscip. Rev. RNA* **2020**, *11*, No. e1587.
(37) Tokunaga, M.; Imamoto, N.; Sakata-Sogawa, K. Highly inclined thin illumination enables clear single-molecule imaging in cells. *Nat. Methods* **2008**, *5*, 159−161.
(38) Lelek, M.; Gyparaki, M. T.; Beliu, G.; Schueder, F.; Griffié, J.; Manley, S.; Jungmann, R.; Sauer, M.; Lakadamyali, M.; Zimmer, C. Single-molecule localization microscopy. *Nat. Rev. Methods Primers* **2021**, *1*, 1−27.
(39) Dmitrieff, S.; Nédélec, F. ConfocalGN: A minimalistic confocal image generator. *Softwarex* **2017**, *6*, 243−247.
(40) Kucherenko, S.; Albrecht, D.; Saltelli, A. Exploring multi-dimensional spaces: a Comparison of Latin Hypercube and Quasi Monte Carlo Sampling Techniques. 2015, DOI: 10.48550/arXiv.1505.02350 (acccessed 2023/04/03).
(41) Renardy, M.; Joslyn, L. R.; Millar, J. A.; Kirschner, D. E. To Sobol or not to Sobol? The effects of sampling schemes in systems biology applications. *Math Biosci* **2021**, *337*, No. 108593.
(42) Forman-Kay, J. D.; Ditlev, J. A.; Nosella, M. L.; Lee, H. O. What are the distinguishing features and size requirements of biomolecular condensates and their implications for RNA-containing condensates? *RNA* **2022**, *28*, 36−47.
(43) Lin, Y.; Protter, D. S. W.; Rosen, M. K.; Parker, R. Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Mol. Cell* **2015**, *60*, 208−219.
(44) Zhang, B.; Zerubia, J.; Olivo-Marin, J. C. Gaussian approximations of fluorescence microscope point-spread function models. *Appl. Opt.* **2007**, *46*, 1819−1829.
(45) Morisaki, T.; Lyon, K.; DeLuca, K. F.; DeLuca, J. G.; English, B. P.; Zhang, Z.; Lavis, L. D.; Grimm, J. B.; Viswanathan, S.; Looger, L. L.; et al. Real-time quantification of single RNA translation dynamics in living cells. *Science* **2016**, *352*, 1425−1429.
(46) Lyon, K.; Aguilera, L. U.; Morisaki, T.; Munsky, B.; Stasevich, T. J. Live-Cell Single RNA Imaging Reveals Bursts of Translational Frameshifting. *Mol. Cell* **2019**, *75*, 172−183.e9.
(47) Boersma, S.; Khuperkar, D.; Verhagen, B. M. P.; Sonneveld, S.; Grimm, J. B.; Lavis, L. D.; Tanenbaum, M. E. Multi-Color Single-Molecule Imaging Uncovers Extensive Heterogeneity in mRNA Decoding. *Cell* **2019**, *178*, 458−472.e19.
(48) Hoek, T. A.; Khuperkar, D.; Lindeboom, R. G. H.; Sonneveld, S.; Verhagen, B. M. P.; Boersma, S.; Vermeulen, M.; Tanenbaum, M. E. Single-Molecule Imaging Uncovers Rules Governing Nonsense-Mediated mRNA Decay. *Mol. Cell* **2019**, *75*, 324−339.e11.
(49) Wu, J.; Zaccara, S.; Khuperkar, D.; Kim, H.; Tanenbaum, M. E.; Jaffrey, S. R. Live imaging of mRNA using RNA-stabilized fluorogenic proteins. *Nat. Methods* **2019**, *16*, 862−865.
(50) Khuperkar, D.; Hoek, T. A.; Sonneveld, S.; Verhagen, B. M. P.; Boersma, S.; Tanenbaum, M. E. Quantification of mRNA translation in live cells using single-molecule imaging. *Nat. Protoc.* **2020**, *15*, 1371−1398.
(51) Ruijtenberg, S.; Sonneveld, S.; Cui, T. J.; Logister, I.; de Steenwinkel, D.; Xiao, Y.; MacRae, I. J.; Joo, C.; Tanenbaum, M. E. mRNA structural dynamics shape Argonaute-target interactions. *Nat. Struct. Mol. Biol.* **2020**, *27*, 790−801.
(52) Calapez, A.; Rosa, A. A statistical pixel intensity model for segmentation of confocal laser scanning microscopy images. *IEEE Trans. Image Process.* **2010**, *19*, 2408−2418.
(53) Herberich, G.; Windoffer, R.; Leube, R. E.; Aach, T. Signal and Noise Modeling in Confocal Laser Scanning Fluorescence Microscopy. In *Medical Image Computing and Computer-Assisted Intervention − MICCAI* 2012, Berlin: Heidelberg, 2012, 2012; Ayache, N., Delingette, H., Golland, P., Mori, K., Ayache, N., Delingette, H., Golland, P., Mori, K., Eds.; Berlin: Heidelberg, *2012*; pp. 381−388, DOI: 10.1007/978-3-642-33415-3_47.
(54) Otsu, N. A Threshold Selection Method from Gray-Level Histograms. *IEEE Trans. Syst., Man, and Cybern.* **1979**, *9*, 62−66.
(55) Bradski, G. The OpenCV library. *Dr Dobbs J.* **2000**, *25*, 120.
(56) Canny, J. A computational approach to edge detection. *IEEE Trans. Pattern Anal Mach Intell.* **1986**, *PAMI-8*, 679−698.
(57) van der Walt, S.; Schönberger, J. L.; Nunez-Iglesias, J.; Boulogne, F.; Warner, J. D.; Yager, N.; Gouillart, E.; Yu, T. scikit-image: image processing in Python. *PeerJ* **2014**, *2*, No. e453.
(58) Kong, H.; Akakin, H. C.; Sarma, S. E. A generalized Laplacian of Gaussian filter for blob detection and its applications. *IEEE Trans. Cybern.* **2013**, *43*, 1719−1733.
(59) Berg, S.; Kutra, D.; Kroeger, T.; Straehle, C. N.; Kausler, B. X.; Haubold, C.; Schiegg, M.; Ales, J.; Beier, T.; Rudy, M.; et al. ilastik: interactive machine learning for (bio)image analysis. *Nat. Methods* **2019**, *16*, 1226−1232.
(60) Lester, E.; Van Alstyne, M.; McCann, K. L.; Reddy, S.; Cheng, L. Y.; Kuo, J.; Pratt, J.; Parker, R. Cytosolic condensates rich in polyserine define subcellular sites of tau aggregation. *Proc. Natl. Acad. Sci. U. S. A.* **2023**, *120*, No. e2217759120.
(61) Chicco, D.; Jurman, G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics* **2020**, *21*, 6.
(62) Harmon, T. S.; Holehouse, A. S.; Rosen, M. K.; Pappu, R. V. Intrinsically disordered linkers determine the interplay between phase separation and gelation in multivalent proteins. *Elife* **2017**, *6*, No. e30294.
(63) Michieletto, D.; Marenda, M. Rheology and Viscoelasticity of Proteins and Nucleic Acids Condensates. *JACS Au.* **2022**, *2*, 1506−1521.

(64) Berry, J.; Brangwynne, C. P.; Haataja, M. Physical principles of intracellular organization via active and passive phase transitions. *Rep. Prog. Phys.* **2018**, *81*, No. 046601.
(65) Morisaki, T.; Stasevich, T. J. Single-Molecule Imaging of mRNA Interactions with Stress Granules. In *The Integrated Stress Response*, Matějů, D., Chao, J. A., Matějů, D., Chao, J. A. Eds.; Vol. *2428*; 2022; pp. 349−360, DOI: 10.1007/978-1-0716-1975-9_21.
(66) Baggett, D. W.; Medyukhina, A.; Tripathi, S.; Shirnekhi, H. K.; Wu, H.; Pounds, S. B.; Khairy, K.; Kriwacki, R. An Image Analysis Pipeline for Quantifying the Features of Fluorescently-Labeled Biomolecular Condensates in Cells. *Front. Bioinf.* **2022**, *2*, No. 897238.
(67) Lee, D. S. W.; Wingreen, N. S.; Brangwynne, C. P. Chromatin mechanics dictates subdiffusion and coarsening dynamics of embedded condensates. *Nat. Phys.* **2021**, *17*, 531−538.
(68) Klickstein, J. A.; Mukkavalli, S.; Raman, M. AggreCount: an unbiased image analysis tool for identifying and quantifying cellular aggregates in a spatially defined manner. *J. Biol. Chem.* **2020**, *295*, 17672−17683.
(69) Yen, J. C.; Chang, F. J.; Chang, S. A new criterion for automatic multilevel thresholding. *IEEE Trans. Image Process.* **1995**, *4*, 370−378.
(70) Keenen, M. M.; Brown, D.; Brennan, L. D.; Renger, R.; Khoo, H.; Carlson, C. R.; Huang, B.; Grill, S. W.; Narlikar, G. J.; Redding, S. HP1 proteins compact DNA into mechanically and positionally stable phase separated domains. *Elife* **2021**, *10*, No. e64563.
(71) Jain, A. K. *Fundamentals of digital image processing*; 1989.
(72) Lindeberg, T. *Scale-space theory in computer vision*; 2013.
(73) Hampoelz, B.; Schwarz, A.; Ronchi, P.; Bragulat-Teixidor, H.; Tischer, C.; Gaspar, I.; Ephrussi, A.; Schwab, Y.; Beck, M. Nuclear Pores Assemble from Nucleoporin Condensates During Oogenesis. *Cell* **2019**, *179*, 671−686.e17.
(74) Lester, E.; Ooi, F. K.; Bakkar, N.; Ayers, J.; Woerman, A. L.; Wheeler, J.; Bowser, R.; Carlson, G. A.; Prusiner, S. B.; Parker, R. Tau aggregates are RNA-protein assemblies that mislocalize multiple nuclear speckle components. *Neuron* **2021**, *109*, 1675−1691.e9.
(75) Han, D.; Longhini, A. P.; Zhang, X.; Hoang, V.; Wilson, M. Z.; Kosik, K. S. Dynamic assembly of the mRNA m6A methyltransferase complex is regulated by METTL3 phase separation. *PLoS Biol.* **2022**, *20*, No. e3001535.
(76) Carrettiero, D. C.; Almeida, M. C.; Longhini, A. P.; Rauch, J. N.; Han, D.; Zhang, X.; Najafi, S.; Gestwicki, J. E.; Kosik, K. S. Stress routes clients to the proteasome via a BAG2 ubiquitin-independent degradation condensate. *Nat. Commun.* **2022**, *13*, 3074.
(77) Yan, V. T.; Narayanan, A.; Wiegand, T.; Jülicher, F.; Grill, S. W. A condensate dynamic instability orchestrates actomyosin cortex activation. *Nature* **2022**, *609*, 597−604.
(78) Geisler, J.; Yan, V. T.; Grill, S.; Narayanan, A. Mass Balance Imaging: A Phase Portrait Analysis for Characterizing Growth Kinetics of Biomolecular Condensates. In *Phase-Separated Biomolecular Condensates: Methods and Protocols*, Zhou, H.-X., Spille, J.-H., Banerjee, P. R., Zhou, H.-X., Spille, J.-H., Banerjee, P. R. Eds.; Methods in Molecular Biology, 2023; pp. 413−424, DOI: 10.1007/978-1-0716-2663-4_21.
(79) Nkoula, S. N.; Velez-Aguilera, G.; Ossareh-Nazari, B.; Hove, L. V.; Ayuso, C.; Legros, V.; Chevreux, G.; Thomas, L.; Seydoux, G.; Askjaer, P.; et al. Mechanisms of Nuclear Pore Complex disassembly by the mitotic Polo-Like Kinase 1 (PLK-1) in C. elegans embryos. *bioRxiv* **2023**, *9*, No. eadf7826.
(80) Kar, M.; Dar, F.; Welsh, T. J.; Vogel, L. T.; Kuhnemuth, R.; Majumdar, A.; Krainer, G.; Franzmann, T. M.; Alberti, S.; Seidel, C. A. M.; et al. Phase-separating RNA-binding proteins form heterogeneous distributions of clusters in subsaturated solutions. *Proc. Natl. Acad. Sci. U. S. A.* **2022**, *119*, No. e2202222119.
(81) Seim, I.; Posey, A. E.; Snead, W. T.; Stormo, B. M.; Klotsa, D.; Pappu, R. V.; Gladfelter, A. S. Dilute phase oligomerization can oppose phase separation and modulate material properties of a ribonucleoprotein condensate. *Proc. Natl. Acad. Sci. U. S. A.* **2022**, *119*, No. e2120799119.
(82) Patel, A.; Lee, H. O.; Jawerth, L.; Maharana, S.; Jahnel, M.; Hein, M. Y.; Stoynov, S.; Mahamid, J.; Saha, S.; Franzmann, T. M.; et al. A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell* **2015**, *162*, 1066−1077.
(83) Lehmussola, A.; Ruusuvuori, P.; Selinummi, J.; Huttunen, H.; Yli-Harja, O. Computational framework for simulating fluorescence microscope images with cell populations. *IEEE Trans. Med. Imaging* **2007**, *26*, 1010−1016.
(84) Stauffer, D.; Solomon, S. Ising, Schelling and self-organising segregation. *Eur. Phys. J. B* **2007**, *57*, 473−479.
(85) Wheeler, J. R.; Matheny, T.; Jain, S.; Abrisch, R.; Parker, R. Distinct stages in stress granule assembly and disassembly. *Elife* **2016**, *5*, No. e18413.