

# Supplementary Information for Spatiotemporal-multiplexed Fourier ptychographic diffraction tomography for high-speed, label-free 3D imaging of live cells

Shun Zhou<sup>1,2,3</sup>, Qian Shen<sup>1,2,3</sup>, Habib Ullah<sup>1,2,3</sup>, Kaiyu Du<sup>1,2,3</sup>, Linpeng Lu<sup>1,2,3</sup>, Hongjun Wu<sup>1,2,3</sup>, Fan Yao<sup>1,2,3</sup>, Jiasong Sun<sup>1,2,3,4</sup>, Dayong Jin<sup>5,\*\*\*</sup>, Qian Chen<sup>3,4,\*\*</sup>, and Chao Zuo<sup>1,2,3,4,\*</sup>

<sup>1</sup>Smart Computational Imaging Laboratory (SCILab), School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210094, China

<sup>2</sup>Smart Computational Imaging Research Institute (SCIRI) of Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210019, China

<sup>3</sup>Jiangsu Key Laboratory of Visual Sensing & Intelligent Perception, Nanjing, Jiangsu Province 210094, China

<sup>4</sup>State key Laboratory of Extreme Environment Optoelectronic Dynamic Measurement Technology and Instrument, Taiyuan, Shanxi Province 030051, China

<sup>5</sup>Institute for Biomedical Materials and Devices (IBMD), Faculty of Science, University of Technology Sydney, Sydney, NSW 2007, Australia

\* zuochao@njust.edu.cn

\*\* chenqian@njust.edu.cn

\*\*\* dayong.jin@uts.edu.au

## ABSTRACT

This document provides supplementary information for “Spatiotemporal-multiplexed Fourier ptychographic diffraction tomography for high-speed, label-free 3D imaging of live cells”. We discuss in detail the implementation of Spatiotemporal-multiplexed Fourier ptychographic diffraction tomography (STM-FPDT) and present more experimental results.

## **Contents**

**Supplementary Note 1. Illumination strategy analysis in STM-FPDT**

**Supplementary Note 2. Verification of the partially coherent forward model of STM-FPDT**

**Supplementary Note 3. Spatiotemporal collaborative RI tomogram reconstruction algorithm of STM-FPDT**

**Supplementary Note 4. Selection of spatial and temporal regularization parameters in STM-FPDT**

**Supplementary Note 5. Experimental evaluation of temporal regularization in STM-FPDT**

**Supplementary Note 6. Clarity evaluation functions for assessing the quality of 3D tomographic imaging**

**Supplementary Note 7. Comparing RI reconstructions with STM-FPDT and aIDT techniques under varied exposure times**

**Supplementary Note 8. Analysis and benchmarking of imaging resolution**

**Supplementary Note 9. Analysis and benchmarking of imaging depth performance**

**Supplementary Note 10. Organelles tracking and 3D cell segmentation algorithm**

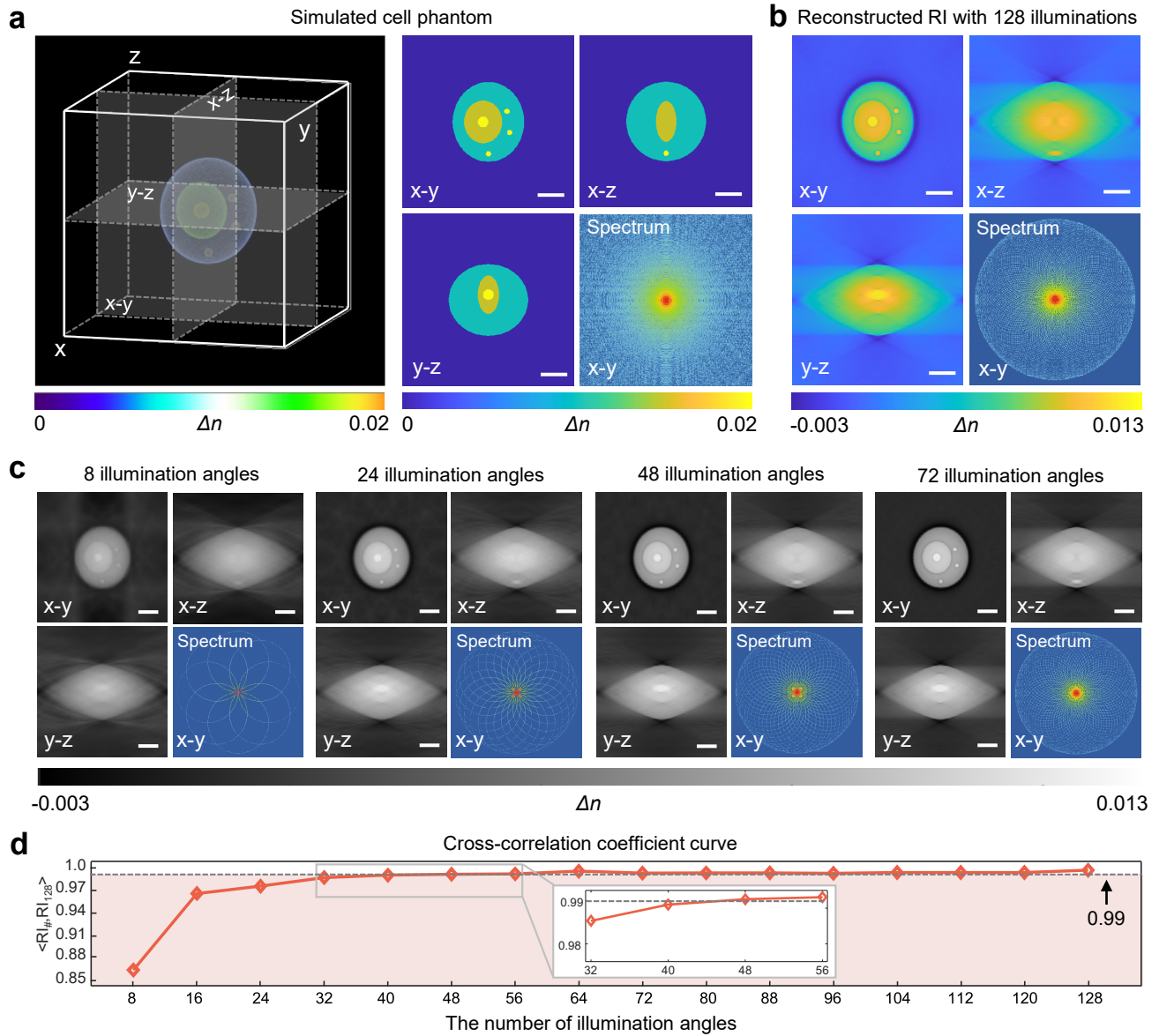
**Supplementary Note 11. Tracing PLC cells migration process with STM-FPDT over a period of two hours**

## Supplementary Note 1. Illumination strategy analysis in STM-FPDT

In Fourier ptychographic diffraction tomography (FPDT) [28], the quality of tomographic imaging is intrinsically related to the number of illumination angles, which determines the extent of accessible object spectral information on the one hand, and affects the convergence during the iterative reconstruction process on the other. An excessive number of illumination angles prolongs the data acquisition time per frame, thereby reducing the imaging speed. Conversely, an insufficient number of illumination angles results in poor due to insufficient spectral information acquisition. Therefore, it is essential to determine the minimum illumination angles required for effective reconstruction, balancing imaging quality and data requisition volume.

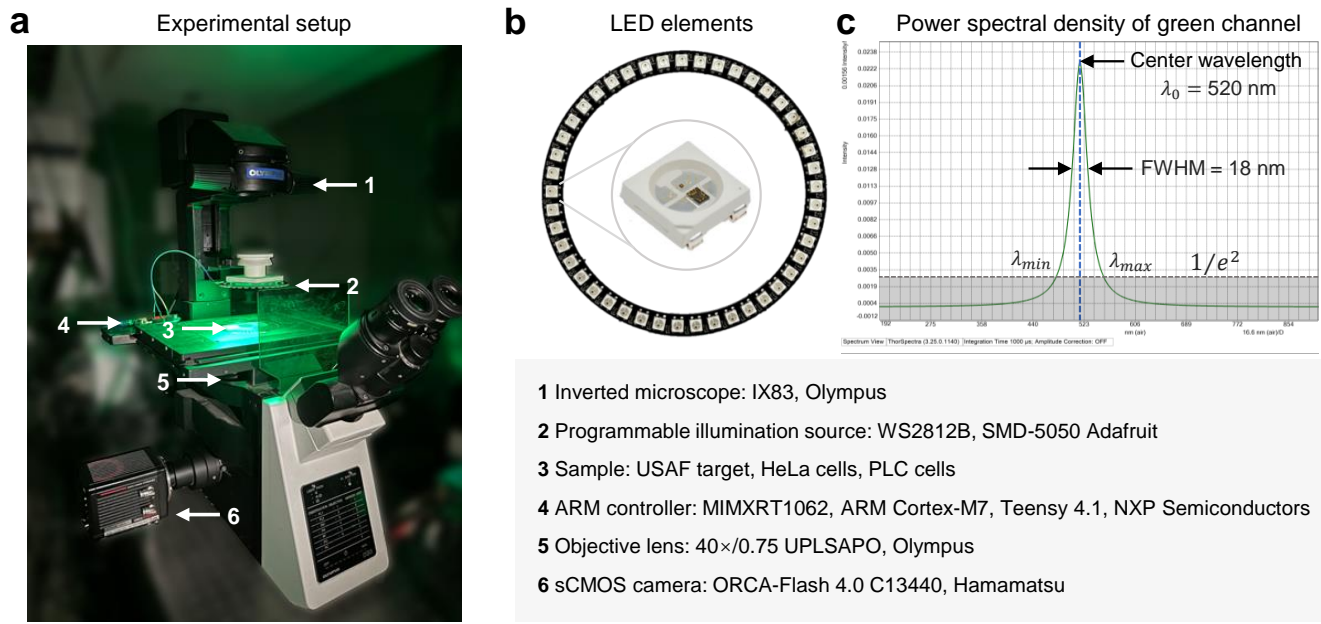
To quantify the relationship between the number of illumination angles and imaging quality, we conducted simulation experiments. Supplementary Fig. S1a illustrates the three-dimensional (3D) rendering of simulated cell phantom, along with refractive index (RI) slices corresponding to x-y, x-z, and y-z planes. Using a  $40\times/0.75$  objective as the detection lens and numerical aperture (NA)-matched annular illumination as the light source, we reconstructed the RI distribution of the cell phantom using FPDT with 128 illumination angles, (Supplementary Fig. S1b). This reconstruction serves as the ground truth for comparative analysis, as 128 illumination angles can access nearly all the information within the theoretical “donut” support domain. The RI distributions and the corresponding x-y plane spectral distributions reconstructed using FPDT with 8, 24, 48, and 72 illumination angles are presented in Supplementary Fig. S1c. Insufficient number angles lead to fuzzy reconstruction results, RI underestimation, and artifacts. As the number of illumination angles increases, the tomographic imaging quality improves progressively. The reconstructed RI distribution with 48 illumination angles is nearly indistinguishable from the ground truth. Further increases in the number of illumination angles result in only marginal improvements in imaging quality. Supplementary Fig. S1d shows the cross-correlation coefficient curve between the reconstructed RI results and the ground truth for different numbers of illumination angles. Using a the cross-correlation coefficient of 0.99 as the threshold, we determined that at least 40 illumination angles are required to achieve high-quality RI tomographic imaging. Considering the existing hardware facilities, we selected a programmable annular LED containing 45 elements as the illumination light source. This choice provides a practical balance between imaging quality and data acquisition efficiency, ensuring robust and reliable tomographic imaging performance. The hardware platform for STM-FPDT is shown in Supplementary Fig. S2, which is constructed based on a commercial inverted microscope equipped with a programmable annular LED illumination source consisting of 45 LED elements. Each element can be controlled independently by an ARM controller.

Under NA-matched LED illumination, we first employ two asymmetric semi-annular coherent illumination patterns. These patterns enable parallel acquisition of the scattering potential spectrum, capturing the scattered field information associated with all 45 LEDs with substantially improved acquisition efficiency. However, although such multiplexing significantly enhances capture efficiency, it inevitably exacerbates the ill-posedness of the inverse problem and increases computational complexity. To stabilize the reconstruction process, we additionally acquire a limited number of independent scattered field measurements under coherent illumination. These measurements provide essential data redundancy that supports stable



**Supplementary Fig. S1. Analysis of minimum number of illumination angles required, based on a simulation sample.** **a**, 3D rendering of simulated cell phantom along with RI slices corresponding to x-y, x-z, and y-z planes. **b**, Reconstructed RI results using FPDT with 128 illumination angles. treated as the ground truth for comparative analysis of image quality under varying numbers of illumination angles. **c**, Reconstructed RI results and the corresponding x-y plane spectral distributions obtained using FPDT with different number of illumination angles. **d**, Cross-correlation coefficient curve between the reconstructed RI results under different number of illumination angles and the ground truth in (b). Scale bars: 5  $\mu\text{m}$ .

iterative reconstruction, while still minimizing the total number of acquired images. To determine the minimum number of coherent illumination measurements required, we performed a simulation study. As shown in Supplementary Fig. S3, we compared reconstruction results obtained by introducing 3, 6, 9, and



**Supplementary Fig. S2. Hardware platform for STM-FPDT.** **a**, A photograph of the STM-FPDT platform. **b**, Photos of programmable annular LED source and single LED element. **c**, Power spectral density of green channel.

12 coherent illumination measurements to provide additional data redundancy. The results indicate that with 9 independent coherent illuminations, STM-FPDT enables high-fidelity RI reconstruction. When fewer than 9 coherent measurements are used, residual cross-talk artifacts inevitably appear due to incomplete de-multiplexing. Balancing data efficiency and reconstruction quality, we therefore adopted 9 additional coherent illumination measurements in the final STM-FPDT implementation.

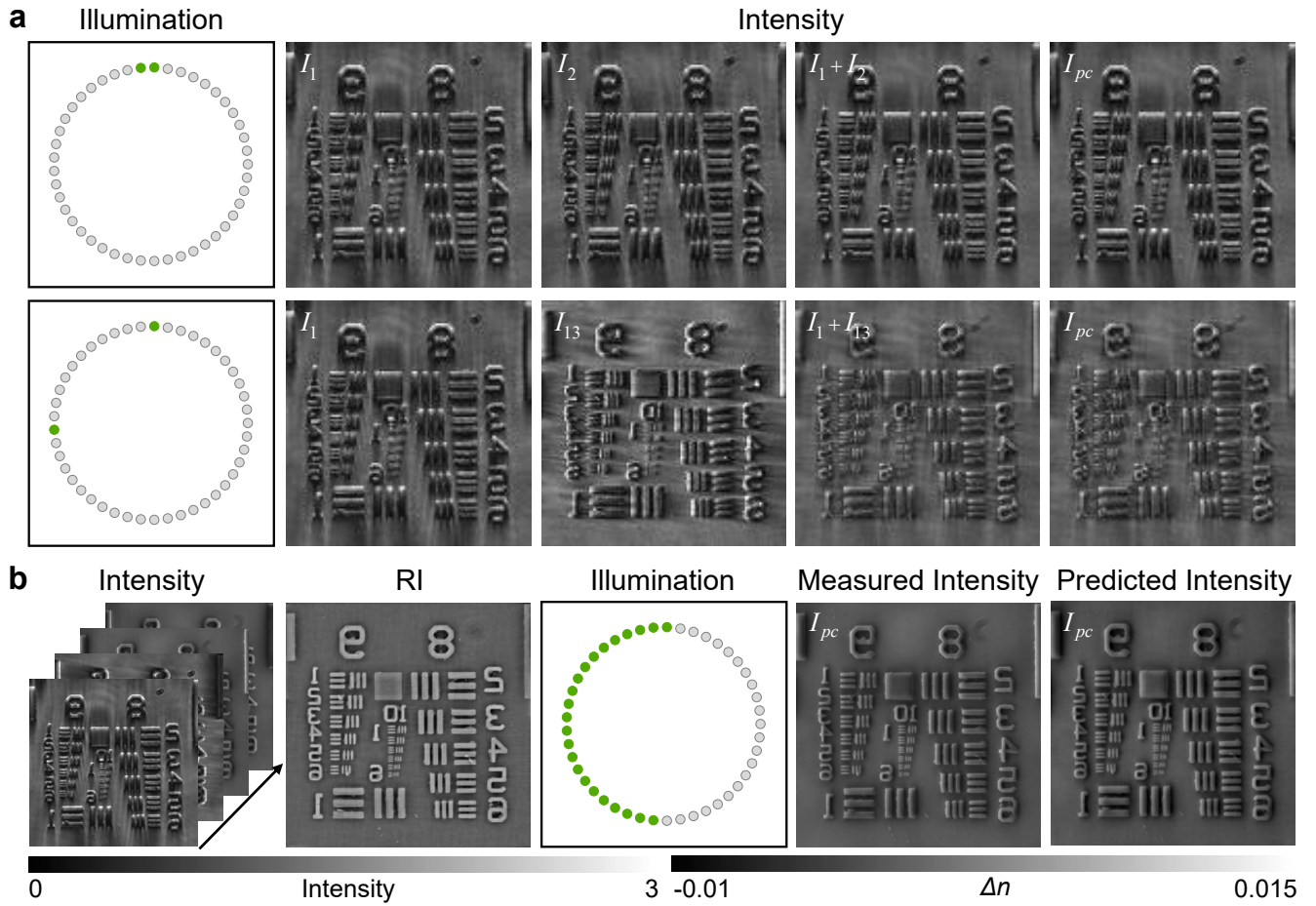
Finally, a revolving illumination strategy was used, which enables independent acquisition of scattering fields under each illumination angle over extended periods, thereby alleviating the burden of field de-multiplexing during reconstruction and ensuring stable, long-term imaging performance. With the 45 LED elements shown in Supplementary Fig. S2b encoded counterclockwise from 1 to 45, Supplementary Table S1 details the illumination groups and their corresponding timing sequence for the 11 images captured in each frame.

**Supplementary Table S1.** Frame indexing scheme under coherent and partially coherent illumination.

Frame number	Coherent illumination									Partially coherent illumination	
	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11
Frame: 1	1	6	11	16	21	26	31	36	41	1~23	24~45
Frame: 2	2	7	12	17	22	27	32	37	42	6~28	29~45, 1~5
Frame: 3	3	8	13	18	23	28	33	38	43	11~33	34~45, 1~10
Frame: 4	4	9	14	19	24	29	34	39	44	15~37	38~45, 1~14
Frame: 5	5	10	15	20	25	30	35	40	45	21~43	44~45, 1~20







**Supplementary Fig. S4. Verification of the partially coherent forward model for STM-FPDT using USAF resolution target.** **a**, Comparison between the summed intensity images obtained by sequentially illuminating each LED individually and the intensity image acquired when the corresponding LEDs are illuminated simultaneously. **b**, Comparison between the experimentally measured asymmetric partially coherent illumination intensity and the predicted intensity during reconstruction process.

### Supplementary Note 3. Spatiotemporal collaborative RI tomogram reconstruction algorithm of STM-FPDT

We developed a spatiotemporal collaborative tomographic reconstruction algorithm to solve the inverse scattering problem and enable high-quality 3D RI imaging. The proposed framework combines a sliding window protocol with a nonlinear global optimization scheme based on the difference map algorithm [43]. This design explicitly exploits the spatiotemporal continuity of dynamic samples to improve convergence robustness and reconstruction stability under highly compressed measurements. In STM-FPDT, the reconstruction is formulated as an optimization problem that estimates the scattering potential (or equivalently, the RI distribution) by minimizing a physically consistent cost function subject to intensity and regularization constraints. The sliding window protocol provides an informed initialization for each time frame by propagating the reconstructed scattering potential from the previous frame, thereby

accelerating convergence and mitigating divergence associated with poor initial estimates. To further stabilize the inverse problem solution, spatiotemporal regularization is incorporated to exploit inter-frame redundancy while suppressing noise and artifacts. Together with sufficient measurement redundancy provided by the hybrid coherent/partially coherent illumination scheme, this formulation ensures stable and efficient convergence of the nonlinear difference map iterations. The pseudo-code for the complete reconstruction workflow of STM-FPDT is outlined in **Algorithm 1**, **Algorithm 2**, and **Algorithm 3**.

---

**Algorithm 1** Forward intensity estimation

---

**Input:** Scattering potential spectrum of the previous frame  $\hat{O}^{pf}(\mathbf{u})$ , current estimated scattering potential spectrum  $\hat{O}(\mathbf{u})$ , incident plane wave  $\{U_{in,i}\}_{i=1}^N$  and plane wave vector  $\{\mathbf{u}_{in,i}\}_{i=1}^N$ ,  $L$  LED elements simultaneously lit under partially coherent illumination.

- 1: Initialization:  $\hat{O}(\mathbf{u}) \leftarrow \hat{O}^{pf}(\mathbf{u})$  ▷ effective initialization
- 2: **for**  $i \leftarrow 1$  to  $N$  **do** ▷ coherent illumination forward model
- 3:    $\hat{U}_{s1,i}(\mathbf{u}_T) \leftarrow \frac{1}{4\pi ju_z} \hat{O}(\mathbf{u} - \mathbf{u}_{in,i}) P_{3D}(\mathbf{u})$
- 4:    $U_{s1,i}(\mathbf{r}_T) \leftarrow \hat{U}_{s1,i}(\mathbf{u}_T)$
- 5:    $I_{c,i} \leftarrow |U_{in,i}(\mathbf{r}_T, \mathbf{u}_{in,i}) \exp[U_{s1,i}(\mathbf{r}_T)/U_{in,i}(\mathbf{r}_T, \mathbf{u}_{in,i})]|^2$
- 6: **end for**
- 7: **for**  $v \leftarrow 1$  to  $2$  **do** ▷ partially coherent illumination forward model
- 8:    $I_{pc,v}(\mathbf{r}_T) \leftarrow 0$
- 9:   **for**  $l \leftarrow 1$  to  $L$  **do**   ( $l \in \{i = 1, 2, \dots, N\}$ )
- 10:      $\hat{U}_{s1,l}(\mathbf{u}_T) \leftarrow \frac{1}{4\pi ju_z} \hat{O}(\mathbf{u} - \mathbf{u}_{in,l}) P_{3D}(\mathbf{u})$
- 11:      $U_{s1,l}(\mathbf{r}_T) \leftarrow \hat{U}_{s1,l}(\mathbf{u}_T)$
- 12:      $I_{c,l} \leftarrow |U_{in,l}(\mathbf{r}_T, \mathbf{u}_{in,l}) \exp[U_{s1,l}(\mathbf{r}_T)/U_{in,l}(\mathbf{r}_T, \mathbf{u}_{in,l})]|^2$
- 13:      $I_{pc,v}(\mathbf{r}_T) \leftarrow I_{pc,v}(\mathbf{r}_T) + I_{c,l}(\mathbf{r}_T)$
- 14:   **end for**
- 15: **end for**

**Return:** Coherent imaging intensity estimation  $\{I_{c,i}(\mathbf{r}_T)\}_{i=1}^N$ , and partially coherent imaging intensity estimation  $\{I_{pc,v}(\mathbf{r}_T)\}_{v=1}^2$ .

---



---

**Algorithm 2** Update the scattering potential based on intensity constraint
 

---

**Input:** Estimated intensity images  $\{I_{c,i}(\mathbf{r}_T)\}_{i=1}^N$ ,  $\{I_{pc,v}(\mathbf{r}_T)\}_{v=1}^2$ , measured intensity images  $\{I_{c,w}^m(\mathbf{r}_T)\}_{w=1}^W$ ,  $\{I_{pc,v}^m\}_{v=1}^2$ , current estimated scattering potential spectrum  $\hat{O}(\mathbf{u})$ , incident plane wave  $\{U_{in,i}\}_{i=1}^N$ , plane wave vector  $\{\mathbf{u}_{in,i}\}_{i=1}^N$ ,  $L$  LED elements simultaneously lit under partially coherent, updated step-size  $\alpha$ , and regularization parameter  $\beta$ .

- 1: **for**  $w \leftarrow 1$  to  $W$  **do** ( $w \in \{i = 1, 2, \dots, N\}$ )  $\triangleright$  intensity constraint under coherent illumination
- 2:    $\hat{U}_{s1,w}(\mathbf{u}_T) \leftarrow \frac{1}{4\pi ju_z} \hat{O}(\mathbf{u} - \mathbf{u}_{in,w}) P_{3D}(\mathbf{u})$
- 3:    $U_{s1,w}(\mathbf{r}_T) \leftarrow \hat{U}_{s1,w}(\mathbf{u}_T)$
- 4:    $U_{s1,w}^{new}(\mathbf{r}_T) \leftarrow U_{in,w}(\mathbf{r}_T, \mathbf{u}_{in,w}) \ln \left( \sqrt{\frac{I_{c,w}^m}{I_{c,w}}} \exp[U_{s1,w}(\mathbf{r}_T, \mathbf{u}_{in,w}) / U_{in,w}(\mathbf{r}_T, \mathbf{u}_{in,w})] \right)$
- 5:    $\hat{O}_w^{new}(\mathbf{u}) \leftarrow U_{s1,w}^{new}(\mathbf{r}_T)$
- 6: **end for**
- 7: **for**  $v \leftarrow 1$  to  $2$  **do**  $\triangleright$  intensity constraint under partially coherent illumination
- 8:   **for**  $l \leftarrow 1$  to  $L$  **do** ( $l \in \{i = 1, 2, \dots, N\}$ )
- 9:      $\hat{U}_{s1,l}(\mathbf{u}_T) \leftarrow \frac{1}{4\pi ju_z} \hat{O}(\mathbf{u} - \mathbf{u}_{in,l}) P_{3D}(\mathbf{u})$
- 10:     $U_{s1,l}(\mathbf{r}_T) \leftarrow \hat{U}_{s1,l}(\mathbf{u}_T)$
- 11:     $I_{c,l}^d(\mathbf{r}_T) \leftarrow |U_{in,l}(\mathbf{r}_T, \mathbf{u}_{in,l}) \exp[U_{s1,l}(\mathbf{r}_T) / U_{in,l}(\mathbf{r}_T, \mathbf{u}_{in,l})]|^2 \frac{I_{pc,v}^m(\mathbf{r}_T)}{I_{pc,v}(\mathbf{r}_T)}$
- 12:     $U_{s1,l}^{new}(\mathbf{r}_T) \leftarrow U_{in,l}(\mathbf{r}_T, \mathbf{u}_{in,l}) \ln \left( \sqrt{\frac{I_{c,l}^d}{I_{c,l}}} \exp[U_{s1,l}(\mathbf{r}_T, \mathbf{u}_{in,l}) / U_{in,l}(\mathbf{r}_T, \mathbf{u}_{in,l})] \right)$
- 13:     $\hat{O}_w^{new}(\mathbf{u}) \leftarrow U_{s1,l}^{new}(\mathbf{r}_T)$
- 14:   **end for**
- 15: **end for**
- 16:  $\hat{O}^{new}(\mathbf{u}) \leftarrow \hat{O}(\mathbf{u}) + \alpha \left( \frac{\sum_{i=1}^N P_{3D}^*(\mathbf{u} + \mathbf{u}_{in,i}) \hat{O}_i^{new}(\mathbf{u})}{\sum_{i=1}^N |P_{3D}(\mathbf{u} + \mathbf{u}_{in,i})|^2 + \beta} - \hat{O}(\mathbf{u}) \right)$   $\triangleright$  update with small step size

**Return:** Updated scattering potential spectrum  $\hat{O}^{new}(\mathbf{u})$ .

---

---

**Algorithm 3** Spatiotemporal continuity constraint and RI tomographic imaging

---

**Input:** Updated scattering potential spectrum  $\hat{O}^{new}(\mathbf{u})$ , measured intensity images  $\{I_{c,w}^m(\mathbf{r}_T)\}_{i=1}^W$ ,  $\{I_{pc,v}^m\}_{v=1}^2$ , convergence threshold of the cost function  $\varepsilon$ , cost function  $C$ , regularization weight  $\tau_{xyz}$ ,  $\tau_t$ ,  $\eta$ , RI in medium  $n_m$ , and wave-number in free space  $k_0$ .

- 1:  $\hat{O}(\mathbf{u}) \leftarrow \hat{O}^{new}(\mathbf{u})$
- 2:  $O(\mathbf{r}) \leftarrow \hat{O}(\mathbf{u})$
- 3:  $\{I_{c,i}(\mathbf{r}_T)\}_{i=1}^N, \{I_{pc,v}(\mathbf{r}_T)\}_{v=1}^2 \leftarrow \text{run Algorithm 1}$
- 4:  $C \leftarrow \frac{1}{2K} \sum_{k=1}^K \|I_k(\mathbf{r}_T) - I_k^m(\mathbf{r}_T)\|_{\ell_2}^2 + \tau_{xyz} \|D_{xyz}(O(\mathbf{r}))\|_{\ell_2} + \tau_t \|D_t(O(\mathbf{r}))\|_{\ell_2}$
- 5: **while**  $C \geq \varepsilon$  **do** ▷ spatiotemporal regularization
- 6:    $\hat{O}^{new}(\mathbf{u}) \leftarrow \text{run Algorithm 2}$
- 7:    $O(\mathbf{r}) \leftarrow \hat{O}^{new}(\mathbf{u})$
- 8:    $O_{grad}(\mathbf{r}) \leftarrow O(\mathbf{r})$  calculate using proximal gradient method
- 9:    $O(\mathbf{r}) \leftarrow O(\mathbf{r}) - \eta O_{grad}(\mathbf{r})$
- 10: **end while**
- 11:  $n(\mathbf{r}) \leftarrow \sqrt{\frac{1}{k_0^2} O(\mathbf{r}) + n_m^2}$

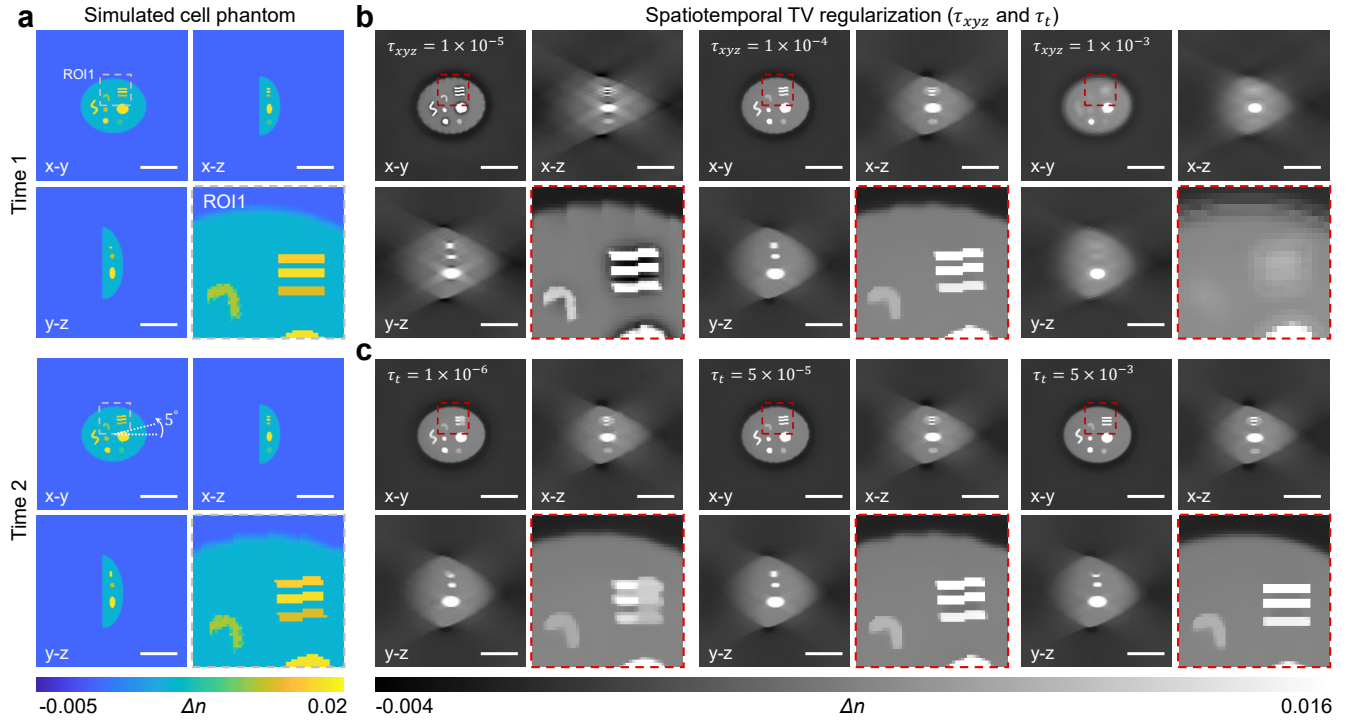
**Return:** Reconstructed 3D RI distribution  $n(\mathbf{r})$ .

---

## Supplementary Note 4. Selection of spatial and temporal regularization parameters in STM-FPDT

STM-FPDT employs a multiplexed illumination strategy to reconstruct a spectrum equivalent to 45 individual illuminations using only 11 measurements. While this multiplexing scheme significantly improves acquisition efficiency, it inevitably increases the ill-posedness of the reconstruction problem and the associated computational burden. Specifically, each illumination component must be computationally de-multiplexed, which can become unstable in the presence of limited measurement numbers and experimental noise. To address this challenge, we introduce spatiotemporal TV regularization into the STM-FPDT reconstruction framework. This regularization exploits the intrinsic piecewise spatial smoothness and temporal consistency of biological structures to stabilize the reconstruction result. Appropriate spatial and temporal regularization is critical for suppressing artifacts and improving reconstruction quality.

Supplementary Fig. S5a shows the simulated cell phantom at Time 1 and its rotated state at Time 2 ( $5^\circ$  rotation about the z axis), used to mimic sample motion. To systematically investigate the influence of the regularization parameters, we use the ideal reconstruction at Time 1 as the prior estimate for Time 2, and analyze the effects of different spatial and temporal regularization strengths on the reconstruction quality at Time 2. With the temporal regularization fixed at  $\tau_t = 5 \times 10^{-5}$ , Supplementary Fig. S5b compares reconstruction results obtained with different spatial regularization parameters  $\tau_{xyz}$ . When  $\tau_{xyz}$  is excessively large, the reconstruction suffers from pronounced oversmoothing, in which fine structural



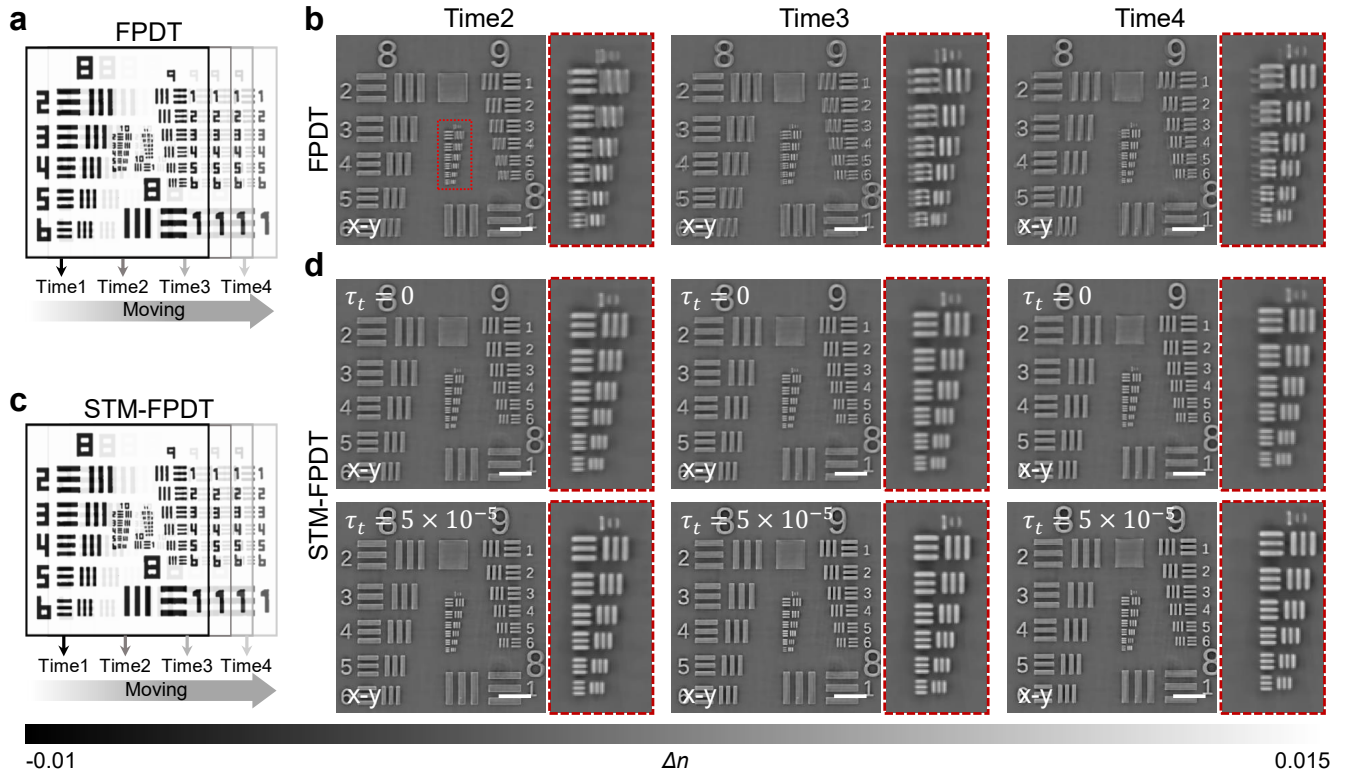
**Supplementary Fig. S5. Selection strategy of spatial and temporal regularization parameters in STM-FPDT based on simulation.** **a**, Simulated cell phantom at Time 1 and its rotated state at Time 2 ( $5^\circ$  rotation about the z axis), used to mimic sample motion. **b**, Comparison of reconstruction results obtained with different spatial regularization parameters  $\tau_{xyz}$ , with the temporal regularization fixed at  $5 \times 10^{-5}$ . **c**, Comparison of reconstruction results obtained with different temporal regularization parameters  $\tau_t$ , with the spatial regularization fixed at  $1 \times 10^{-4}$ . Scale bars: 10  $\mu\text{m}$ .

details are severely suppressed and characteristic cartoon-like artifacts emerge. Conversely, when  $\tau_{xyz}$  is too small, the spatial regularization becomes insufficient to effectively suppress reconstruction artifacts. Similarly, with the spatial regularization fixed at  $\tau_{xyz} = 1 \times 10^{-4}$ , Supplementary Fig. S5c illustrates the effect of different temporal regularization strengths  $\tau_t$  on dynamic reconstructions. An overly large  $\tau_t$  introduces excessive temporal smoothing, manifested as noticeable inter-frame lag, whereas an excessively small  $\tau_t$  provides insufficient temporal constraints to suppress crosstalk artifacts during reconstruction.

Based on these observations,  $\tau_{xyz}$  was set to  $1 \times 10^{-4}$  to effectively suppress spatial artifacts while preserving fine structural features, and  $\tau_t$  was set to  $5 \times 10^{-5}$  to enforce temporal continuity without oversmoothing rapid dynamic changes. Importantly, the reconstruction quality is largely insensitive to small variations of these parameters within a reasonable range, indicating that STM-FPDT achieves a favorable balance between artifact suppression and dynamic fidelity and exhibits a degree of robustness.

## Supplementary Note 5. Experimental evaluation of temporal regularization in STM-FPDT

The fast imaging capability of STM-FPDT arises from compressing the scattered field information from 45 illumination angles into only 11 measurements, allowing a volumetric imaging rate of up to 5 Hz. The introduction of the sliding window protocol and the regularization strategy is aimed at addressing the ill-posed nature of the reconstruction problem and the computational burden caused by the multiplexing strategy. To experimentally evaluate the robustness of STM-FPDT against sample motion and to assess the role of temporal regularization, controlled motion experiments were performed using a customized resolution target (Supplementary Fig. S6). Supplementary Figs. S6a and S6c schematically illustrate the motion control protocol. For both conventional FPDT and STM-FPDT, intensity measurements were



**Supplementary Fig. S6. Experimental evaluation of temporal regularization in STM-FPDT using a customized resolution target.** **a**, Schematic of the controlled motion protocol for the resolution target during conventional FPDT data acquisition. At Time 1, the target remains stationary. **b**, Reconstruction RI distributions at different time points using conventional FPDT, showing pronounced motion-induced artifacts due to prolonged acquisition time. **c**, Schematic of the controlled motion protocol for STM-FPDT acquisition under identical conditions. **d**, STM-FPDT reconstruction results at different time points without temporal regularization ( $\tau_t = 0$ , top row) and with temporal regularization ( $\tau_t = 5 \times 10^{-5}$ , bottom row), demonstrating improved fidelity and reduced motion artifacts when temporal regularization is applied. Scale bars: 10  $\mu\text{m}$ .

acquired at four discrete time points during continuous target motion, with the target remaining stationary at the initial time point (Time 1). Due to the prolonged acquisition time required by conventional FPDT, severe motion-induced artifacts are observed in the reconstructed RI distributions, leading to pronounced degradation of fine structural details (Supplementary Fig. S6b). In contrast, STM-FPDT benefits from its rapid acquisition strategy and exhibits substantially reduced motion artifacts across successive time frames. When temporal regularization is disabled ( $\tau_t = 0$ ), STM-FPDT reconstructions show slight high-frequency blurring, which are attributed to incomplete de-multiplexing (Supplementary Fig. S6d). By incorporating temporal regularization ( $\tau_t = 5 \times 10^{-5}$ ), inter-frame redundancy is effectively exploited, resulting in improved reconstruction fidelity and enhanced robustness against motion-induced inconsistencies.

## Supplementary Note 6. Clarity evaluation functions for assessing the quality of 3D tomographic imaging

In the quality assessment of diffraction tomography, image clarity stands as a vital metric. Here, three clarity evaluation functions are invoked to assess the quality of 3D tomographic imaging, which are gradient of energy (EOG), discrete Fourier transform (DFT) and information entropy.

### 1. Gradient of energy

Sharp images exhibit greater clarity compared to blurred ones, with edge pixels showing significant variations in grey values, resulting in larger gradient values. The EOG function quantifies the edge information variation in an image by computing its first-order derivative, thereby reflecting the image's clarity. Utilizing EOG, the clarity of 3D RI tomographic imaging can be expressed as follows:

$$D_{EOG} = \sum \left( [\nabla_x n(\mathbf{r})]^2 + [\nabla_y n(\mathbf{r})]^2 + [\nabla_z n(\mathbf{r})]^2 \right) \quad (S1)$$

where  $\mathbf{r} = (\mathbf{r}_T, z) = (x, y, z)$  is a short-hand notation for the 3D spatial coordinate.  $\nabla_x$ ,  $\nabla_y$ , and  $\nabla_z$  respectively represent the gradient operators along the  $x$ ,  $y$ , and  $z$  directions.

### 2. Discrete Fourier transform

As a blurred image is akin to a clear image subjected to low-pass filtering, heightened presence of high-frequency components indicates enhanced clarity. By deriving the spectral distribution of the image via DFT, emphasis is placed on high-frequency components, and the weighted sum of spectral components is computed to delineate image clarity. Leveraging the DFT, the clarity of 3D RI tomographic imaging can be articulated as follows:

$$D_{DCT} = \frac{1}{V} \sum \sqrt{u_x^2 + u_y^2 + u_z^2} (\mathcal{F} \{n(\mathbf{r})\}) \quad (S2)$$

where  $V$  is total number of pixels for the 3D RI matrix.  $u_x$ ,  $u_y$ , and  $u_z$  respectively represent the spatial frequency coordinates along the  $x$ ,  $y$ , and  $z$  directions.  $\mathcal{F} \{ \cdot \}$  denotes 3D discrete Fourier transform of functions.

### 3. Entropy

In Shannon's information theory, entropy serves as a common metric for depicting information richness. In a clear image, the distribution of pixel values spans wide intervals and exhibits substantial differences, resulting in a higher entropy value compared to a blurred image. Drawing upon Shannon's Entropy Model, the clarity of 3D RI tomographic imaging can be characterized as follows:

$$D_{Entropy} = -\frac{1}{V} \sum_{q=1}^Q P_q(n(\mathbf{r})) \log \left[ \frac{P_q(n(\mathbf{r}))}{V} \right] \quad (S3)$$

where  $P_q(\cdot)$  is a function used to counts the number of pixels in 3D space that satisfy  $n(\mathbf{r}) = q$ , and  $Q$  is the maximum RI value of  $n(\mathbf{r})$ .

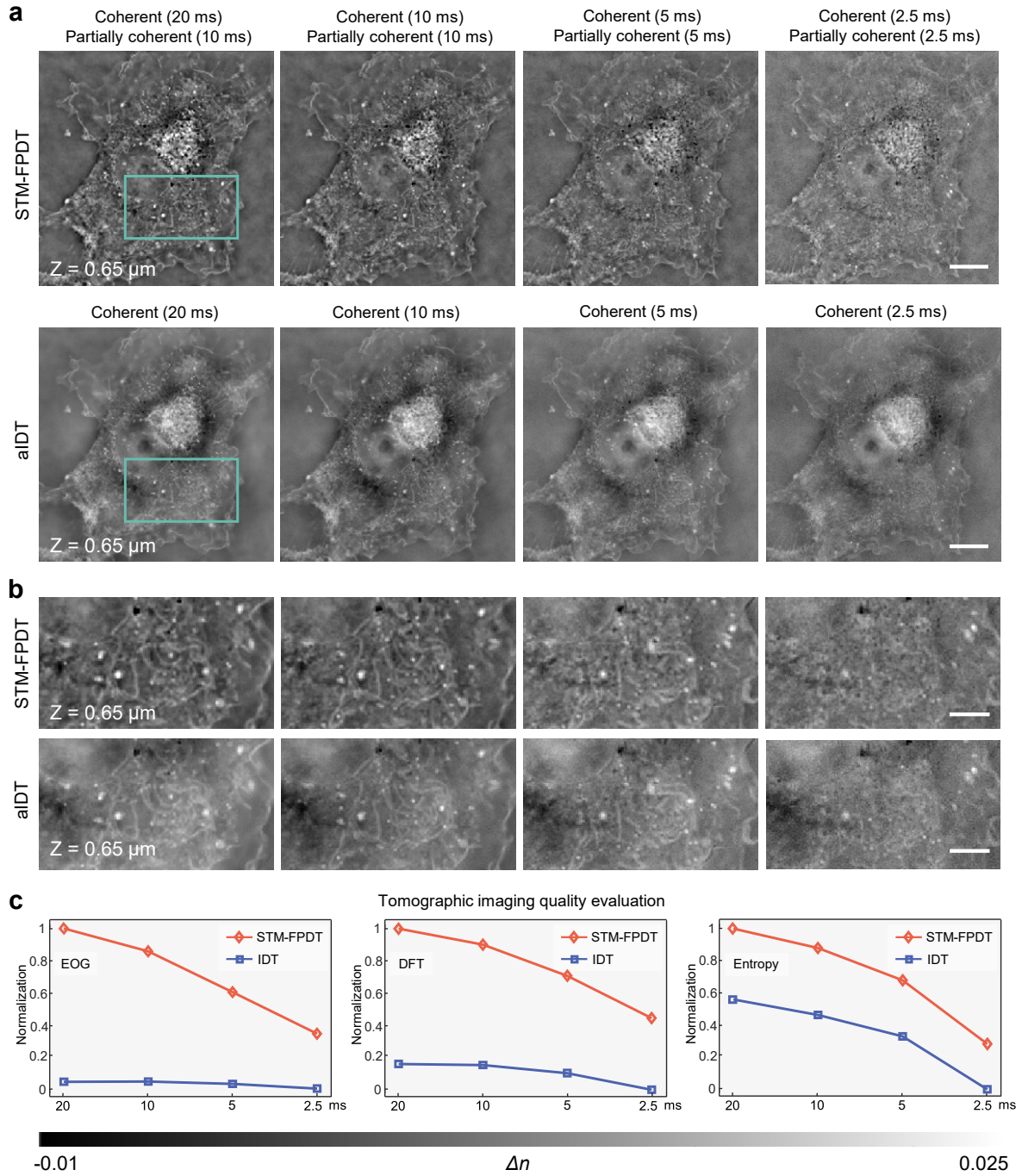
## Supplementary Note 7. Comparing RI reconstructions with STM-FPDT and aIDT techniques under varied exposure times

To illustrate the advantages offered by STM-FPDT, we conducted a comparative analysis of HeLa cell RI reconstructions using STM-FPDT and aIDT [48] techniques with varied exposure times. The process began by acquiring intensity image stacks using both techniques at different exposure durations. Subsequently, these exposure time-varied intensity image stacks were utilized to reconstruct the RI distribution for comparison.

Our analysis revealed that minimizing exposure time can significantly enhance the rate of volumetric imaging. For instance, the RI results, illustrated in Supplementary Fig. S7a, demonstrate that the exposure time for coherent illumination can effectively be reduced from 20 ms to 10 ms without significantly degrading imaging quality based on existing hardware. However, further reduction in exposure time in partially readout mode of camera leads to deterioration of image quality, with subcellular features becoming obscured by noise (Supplementary Fig. S7b). With the adoption of higher power LEDs, we are optimistic that the volumetric imaging rate of STM-FPDT under local field of view (FOV) could achieve speeds up to 30Hz.

Furthermore, compared to the results reconstructed by the aIDT technique, STM-FPDT yields higher-quality RI tomograms characterized by uniform backgrounds and clear subcellular features. This improvement is attributed to the approximation assumption of intensity [29] used in the slice-wise deconvolution reconstruction process of aIDT, and the results are easily vulnerable to noise level and regularization parameters. Inappropriately chosen parameters often result in low-frequency cloud artifacts or the suppression of high-frequency details. In contrast, as an iterative solution strategy based on intensity self-consistency, STM-FPDT relies on fewer assumptions and is less susceptible to regularization parameters. The three clarity evaluation functions introduced in Supplementary Note 2 are utilized to quantitatively access the 3D imaging quality under varying conditions, as depicted in Fig. S7c, indicating that the imaging quality achieved by STM-FPDT consistently surpasses that of IDT across all evaluation functions. Consequently, we assert that STM-FPDT presents significant advantages in imaging quality and holds potential for future applications.

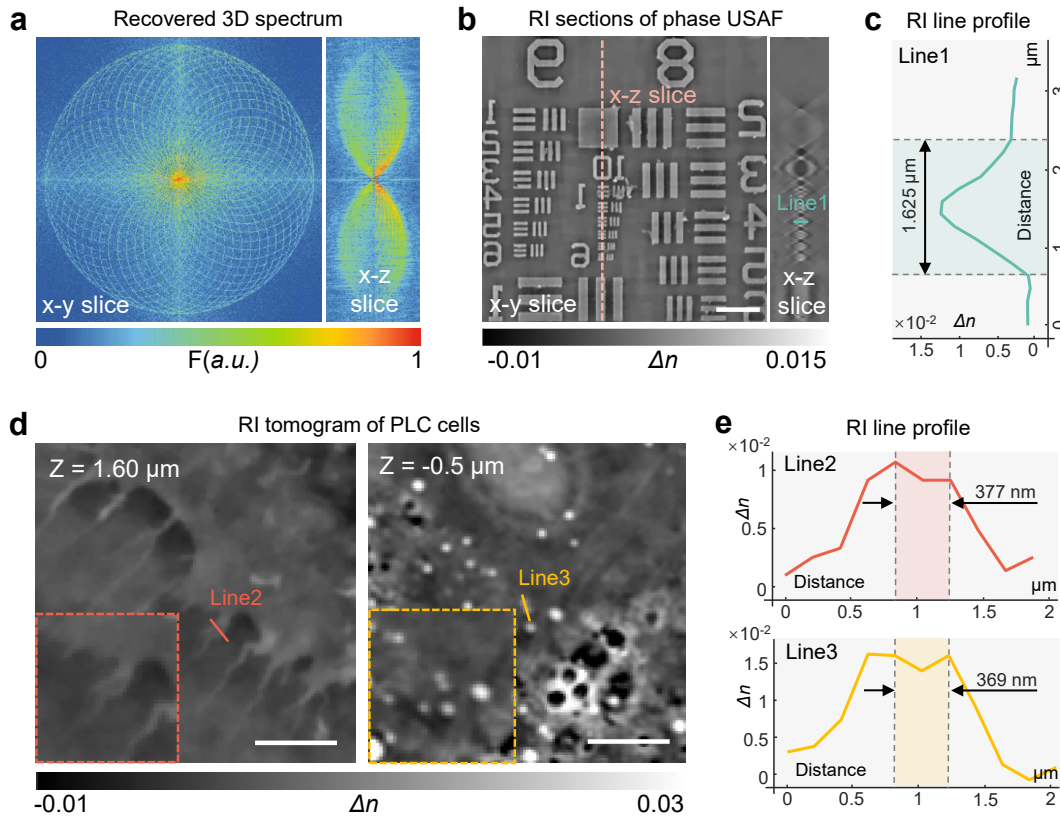




**Supplementary Fig. S7. Comparing RI reconstructions of STM-FPDT and aIDT techniques. a,** Reconstructed RI tomograms of HeLa cells using STM-FPDT and aIDT under varied exposure time. The exposure time under coherent illumination is systematically decreased from 20 ms to 2.5 ms, while under partially coherent illumination, it is gradually reduced from 10 ms to 2.5 ms. **b,** The subregions containing mitochondria within the green rectangular box in (a) are enlarged to highlight the reconstructed subcellular feature. **c,** Three clarity evaluation functions are invoked to assess the quality of 3D tomographic imaging of two methods at different exposure times. Scale bars: (a) 10  $\mu\text{m}$ , and (b) 5  $\mu\text{m}$ .

## Supplementary Note 8. Analysis and benchmarking of imaging resolution

To demonstrate the proposed method's performance, we evaluate our technique's resolution limit by performing STM-FPDT imaging on both phase United States Air Force (USAF) target and subcellular structures within PLC cells. The highest accessible frequency of the reconstructed scattering potential spectrum determines the tomographic imaging resolution limit. The tomographic imaging resolution limit is determined by the highest accessible frequency of the reconstructed scattering potential spectrum [37]. The lateral and axial sections of the recovered 3D RI spectrum of the phase USAF without employing total variation regularization are illustrated in Supplementary Fig. S8a. Based on the experimental parameters



**Supplementary Fig. S8. Resolution benchmarking of STM-FPDT with phase USAF and PLC cells.** **a**, Lateral and axial sections of the recovered high-resolution 3D RI spectrum of phase USAF. **b**, Lateral and axial sections of the reconstructed RI result of phase USAF. **c**, RI line profile corresponding to the marker in **(b)** reveals that the axial width of bar is 1.625  $\mu\text{m}$ . **d**, Two RI tomograms of PLC cells demonstrating the high-resolution imaging of subcellular details. The illustrations at the bottom left enlarge the filopodia and lipid droplets structures, providing a clearer view. **e**, RI line profiles across filopodia and lipid droplets corresponding to the markers in **(d)** validate that the full-width lateral resolution of STM-FPDT can reach 369 nm. Scale bars: **(b)** 10  $\mu\text{m}$ , and **(d)** 5  $\mu\text{m}$ .

provided in the main text and using the resolution definition formula

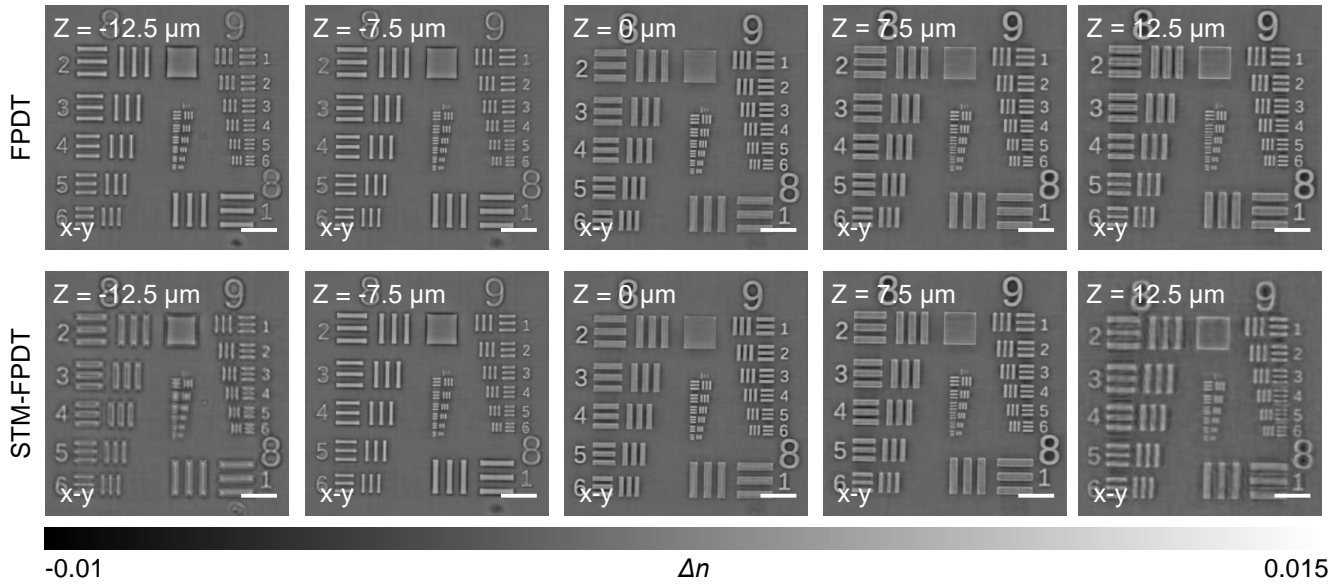
$$\Delta_{x,y} = \frac{\lambda}{NA_{ill} + NA_{obj}}, \quad \Delta_z = \frac{\lambda}{1 - \sqrt{1 - NA_{ill}^2}}, \quad (S4)$$

we calculate that the theoretical full-width lateral resolution is 347 nm, and axial resolution is 1.54  $\mu\text{m}$ , respectively. Supplementary Fig. S8b illustrates the reconstructed RI result at the in-focus plane as well as the axial section corresponding to the position of the pink dashed line. Since the theoretical lateral resolution of the STM-FPDT is about 347 nm, far beyond the minimum bar in the phase target, only element 6 of group 10 is clearly visible despite some damage to the bar at this position. Furthermore, as a single-layer thin object, the axial imaging width of the phase resolution target can be used to benchmark the axial resolution limit. Analysis of the line profile width in Supplementary Fig. S8c confirms the full-width axial resolution of about 1.625  $\mu\text{m}$ , which closely matches the theoretical prediction of 1.54  $\mu\text{m}$ . This slight discrepancy arises from the missing cone issue in the direction near the low zero-frequency region [47], but the regularization alleviates this problem well.

Additionally, the lateral resolution of STM-FPDT is further assessed by the analysis of imaging results of subcellular structures within PCL cells, as depicted in Supplementary Fig. S8d. Line profiles across filopodia and lipid droplets corresponding to the markers in Supplementary Fig. S8d validate a near-diffraction-limited full-width lateral resolution of 369 nm, closely matching with the theoretical prediction of 347 nm, as demonstrated in Supplementary Fig. S8e.

## Supplementary Note 9. Analysis and benchmarking of imaging depth performance

As a focused-imaging-based, non-interferometric optical diffraction tomography technique, the imaging depth of STM-FPDT is governed by several fundamental physical factors. First, the finite spatial and temporal coherence of the illumination source intrinsically limits the achievable imaging depth. The LED source employed in this work is not an ideal coherent source. Its finite spectral bandwidth ( $\sim 10$  nm) and extended emission area ( $\sim 1$  mm) substantially reduce the coherence length. Consequently, scattered fields originating from defocused planes become increasingly blurred at the detector plane, leading to rapid attenuation of high spatial frequency information and, ultimately, a reduced effective imaging depth. Second, under tilted or partially coherent illumination, sample structures located outside the focal plane generate scattered fields that are spatially spread over a wide region at the image plane. When this spread exceeds the detector FOV, a portion of the scattered signal is inevitably lost, preventing accurate encoding and reconstruction of out-of-focus structures. Therefore, the maximum illumination angle and the detector field of view jointly impose an additional constraint on the effective axial imaging range. Third, strong multiple scattering within the sample further limits the achievable imaging depth of STM-FPDT. The current forward model is based on the first-order Rytov approximation, which assumes weak scattering and restricts the optical thickness of the sample. While the coherence-related and FOV limitations may be alleviated at the hardware level by employing illumination sources with higher coherence (at the cost



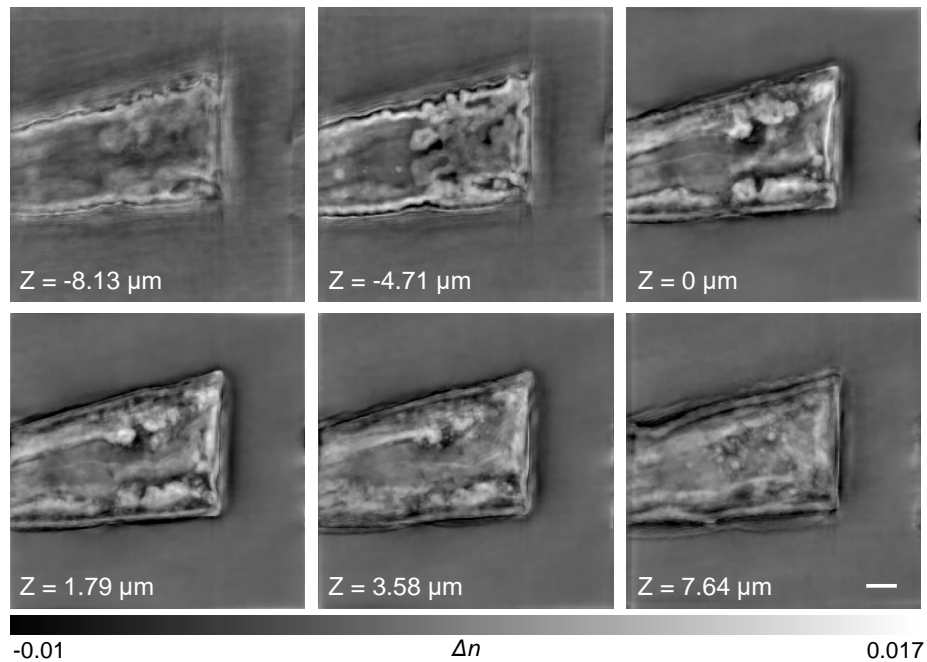
**Supplementary Fig. S9. Quantitative evaluation of axial imaging depth in STM-FPDT and conventional FPDT using defocus measurements.** Reconstruction results of a resolution target placed at different defocused planes for conventional FPDT and STM-FPDT. Scale bars: 10  $\mu\text{m}$ .

of increased speckle noise) or detectors with a larger FOV, overcoming the limitation imposed by strong scattering requires algorithmic advances. In particular, more accurate forward models that account for multiple scattering, such as beam propagation methods, Born series expansions, or data-driven forward modeling using deep learning, together with corresponding inverse solvers, are necessary. Developing such physically faithful and computationally efficient reconstruction frameworks remains an active and challenging research direction.

Notably, the partially coherent illumination adopted in STM-FPDT further shortens the coherence length and introduces an additional coherence gating effect along the axial direction. As a result, scattered fields can only be effectively recorded within a limited axial range. When sample structures extend beyond this range, scattered fields from defocused planes rapidly decay and are easily overwhelmed by noise, preventing their reliable encoding in the measured intensities. Consequently, these out-of-range scattered components cannot be effectively utilized in the subsequent inverse reconstruction. This constraint leads to a moderately reduced effective imaging depth for STM-FPDT compared with conventional FPDT based on fully coherent measurements.

To quantitatively evaluate the practical imaging depth, we placed the resolution target at different defocused planes ( $-12.5\mu\text{m}$ ,  $-7.5\mu\text{m}$ ,  $0\mu\text{m}$ ,  $7.5\mu\text{m}$ , and  $12.5\mu\text{m}$ ) and compared the reconstruction results obtained using conventional FPDT and STM-FPDT. As shown in Supplementary Fig. S9, conventional FPDT achieve effective reconstruction within a range of approximately  $\pm 12.5\mu\text{m}$ , whereas the effective reconstruction range of STM-FPDT is limited to approximately  $\pm 7.5\mu\text{m}$ . Beyond this range, the scattered field information becomes increasingly difficult to record under partially coherent illumination, resulting in degraded reconstruction quality at  $\pm 12.5\mu\text{m}$ . In addition, we experimentally characterized a 3D printed *C. elegans*-like sample to further assess the axial imaging capability of STM-FPDT. As shown in



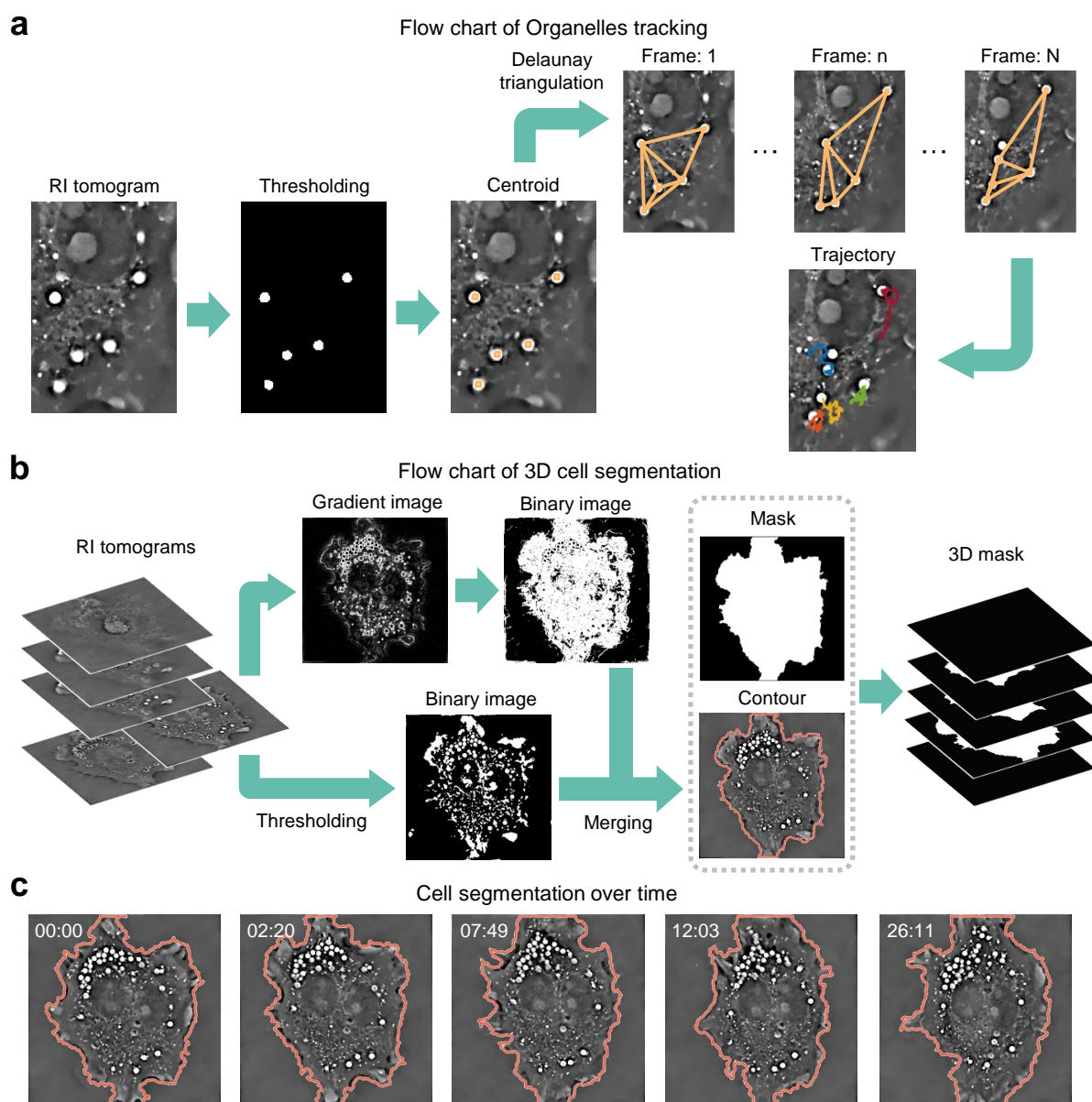


**Supplementary Fig. S10. Imaging depth benchmarking of STM-FPDT using 3D printed sample.** Reconstructed RI slices at different depths demonstrate relatively reliable reconstruction of axial structures within an imaging depth of 15  $\mu\text{m}$ . Scale bars: 20  $\mu\text{m}$ .

Supplementary Fig. S10, reconstructed RI slices at different axial depths demonstrate relatively reliable reconstruction of axial structures within an imaging depth of 15  $\mu\text{m}$ . This imaging depth is sufficient for most biological cell analysis applications, as adherent cells typically exhibit axial dimensions on the order of  $\sim 10 \mu\text{m}$ . For applications requiring tomographic imaging beyond this depth range, axial scanning could be incorporated to extend the effective depth of field and enlarge the measurement range.

## Supplementary Note 10. Organelles tracking and 3D cell segmentation algorithm

A simple yet effective method is employed for the precise localization and tracking of organelles during the analysis biological processes. The flow chart of organelles tracking with lipid droplets as an example is shown in Supplementary Fig. S11a. A thresholding scheme is utilized to robustly segment and extract them based on the RI disparities between lipid droplets, other organelles, and the surrounding medium. The integration of thresholding with an extended maxima transform facilitates the determination of the centroid of each lipid droplet. Subsequently, the motion trajectories of these lipid droplets are established through a tracking framework that evaluates the structural relationships, particularly the topological connections, among neighboring units in consecutive frames. Delaunay triangulation is employed to construct a unique planar graph that is independent of node topology, thereby encoding the spatial relationships between lipid droplets within the cell. The similarity between local Delaunay structures in two consecutive frames is efficiently estimated by accurately modeling the relationships among lipid droplets. By sorting these



**Supplementary Fig. S11. Flow chart of organelles tracking and 3D cell segmentation algorithm.**

**a**, Flow chart of organelles tracking with lipid droplets as an example. **b**, Flow chart of 3D cell segmentation. **c**, The segmentation results of a layer of RI tomograms of 3D cell over time.

similarities, locations exhibiting the highest similarity are identified as corresponding to the same lipid droplet. The trajectory of each lipid droplet is then delineated by associating these positions over time.

To achieve 3D cells segmentation from the background and subsequently measure cell volume, we employ an modified live-cell analysis framework (LAF) tailored for diffraction tomographic imaging (Supplementary Fig. S11b). The key steps of this framework are outlined as follows:

**Step 1:** Selecte a RI tomogram at a specific z-plane and conduct threshold segmentation based on the RI disparities between the cell and the surrounding medium.



**Step 2:** Calculate the gradient of the RI tomogram and utilize Empirical Gradient Threshold (EGT) to segment the gradient image by applying thresholding at various gradient percentile value.

**Step 3:** Merge the threshold segmentation outcomes of **Step 1** and **Step 2** to generate a binary image that retains both the primary structure and fine details of cell edges.

**Step 4:** Implement erosion and expansion techniques to eliminate connected components smaller than the minimum cell area and holes larger than the maximum hole area, with the minimum cell area and maximum hole area as parameters.

**Step 5:** Generate a binary mask comprising a single connected component.

**Step 6:** Repeat **Step 1-5** to compute the mask corresponding to the RI tomogram for each z-plane.

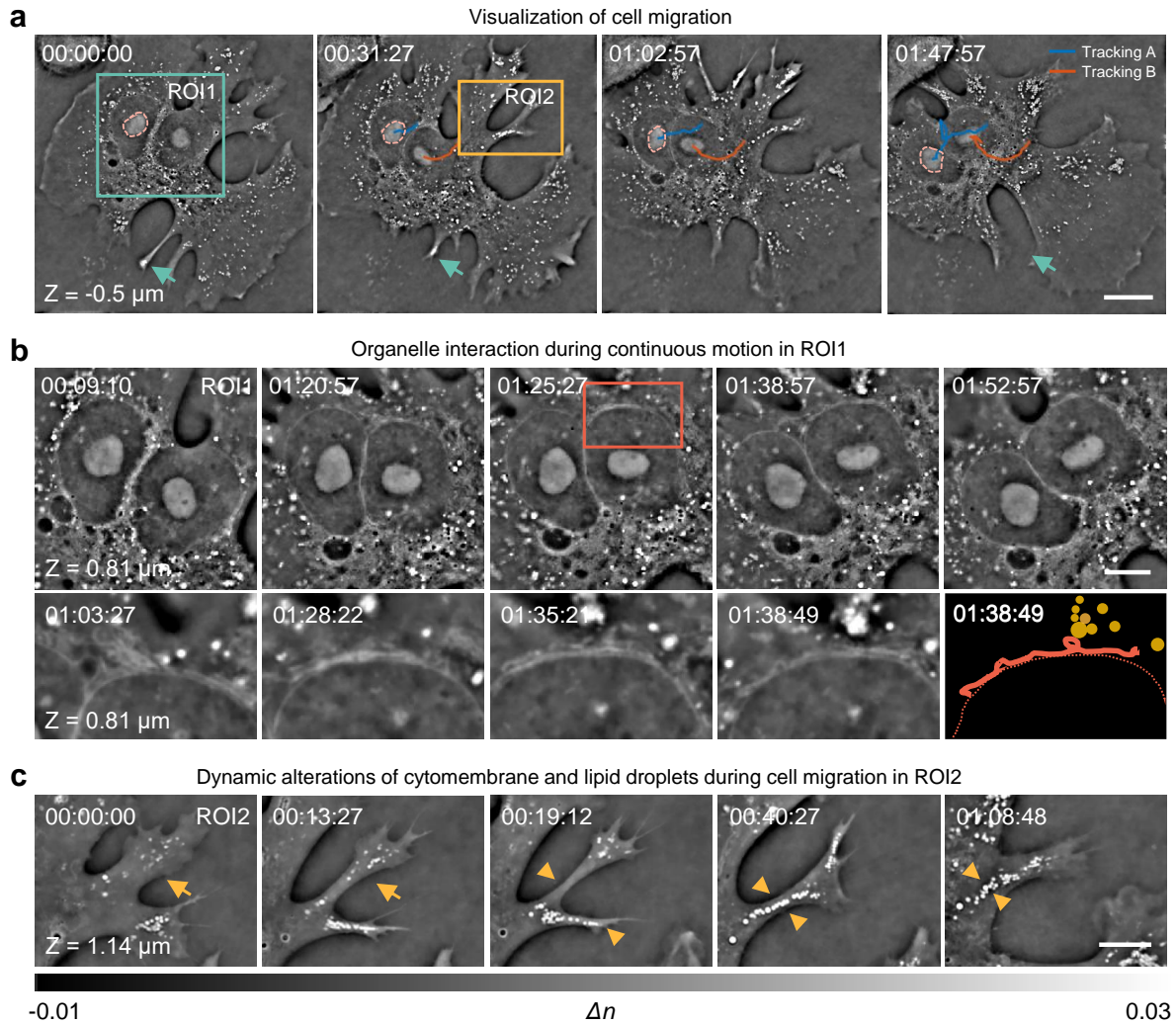
**Step 7:** Calculate the cell volume based on the 3D mask and voxel size.

Supplementary Fig. [S11c](#) presents the segmentation results of a layer of RI tomogram of a cell over time. This methodology ensures accurate segmentation of 3D cell, facilitating precise volumetric measurements essential for various biological analyses.

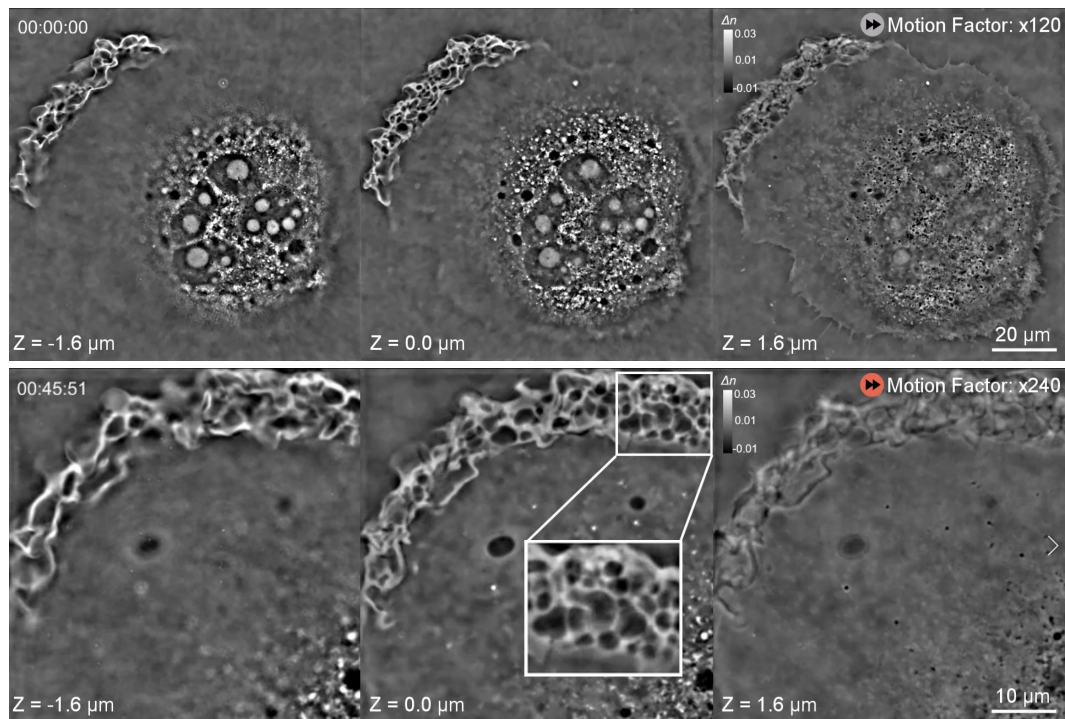
## **Supplementary Note 11. Tracing PLC cells migration process with STM-FPDT over a period of two hours**

Cell migration is a pivotal activity in cell biology, crucial for various physiological and pathological processes such as immune response, tissue repair, and tumor metastasis. A comprehensive understanding and investigation into its mechanisms may reveal the onset and progression of many diseases and provide new insight for future disease treatment and drug development.

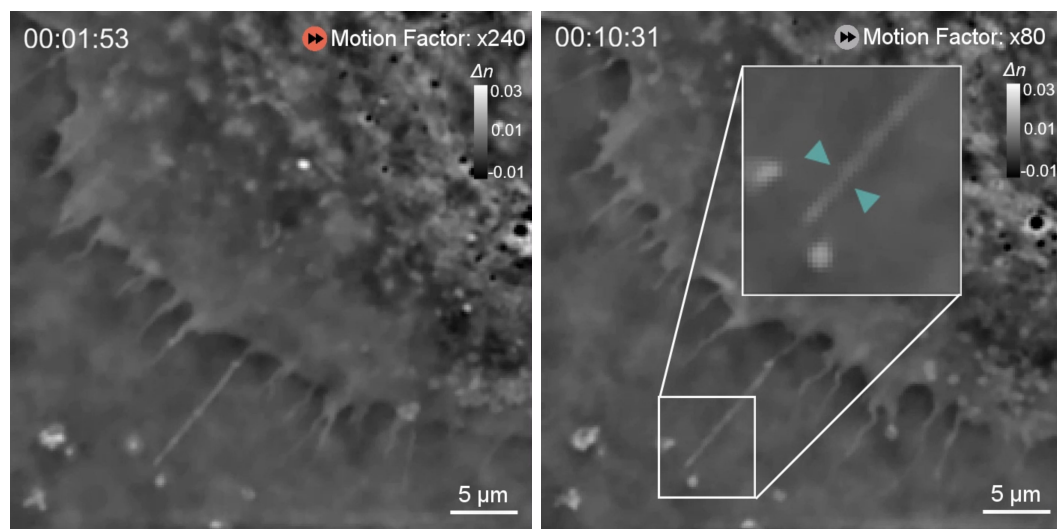
Supplementary Fig. [S12](#) presents both the cellular landscape and subcellular structural changes during cell migration, and the corresponding time-evolution video spanning two hours is provided in Supplementary Video 7. The RI distribution results of  $-0.5\ \mu\text{m}$  z plane at different time points are displayed in Supplementary Fig. [S12a](#), revealing that cell movement is achieved through the continuous adhesion, contraction and depolymerization process between cells and extracellular matrix (green arrow in Supplementary Fig. [S12a](#)). By tracking the nucleolus positions of two cells, we can achieve the precise monitoring of their migratory trajectories. The RI tomograms of ROI1 in Supplementary Fig. [S12a](#) are enlarged in Supplementary Fig. [S12b](#) at different time points, which contains two nuclei with continuous motion. Further, the second row focuses on a small orange box, emphasizing the interactive effects between organelles (including nuclear membranes, mitochondria, and lipid droplets) during the process of cell migration. In particular, the schematic diagram at the time point of 01:38:49 provides a clear depiction of subcellular structure. By continuously monitoring specific regions in PLC living cells over a long period of time using the proposed method, we can clearly observe the production of pseudopods near the cell boundary as well as the behavior of lipid droplets within the cell, which drives the cell in the direction of migration (Supplementary Fig. [S12c](#)).



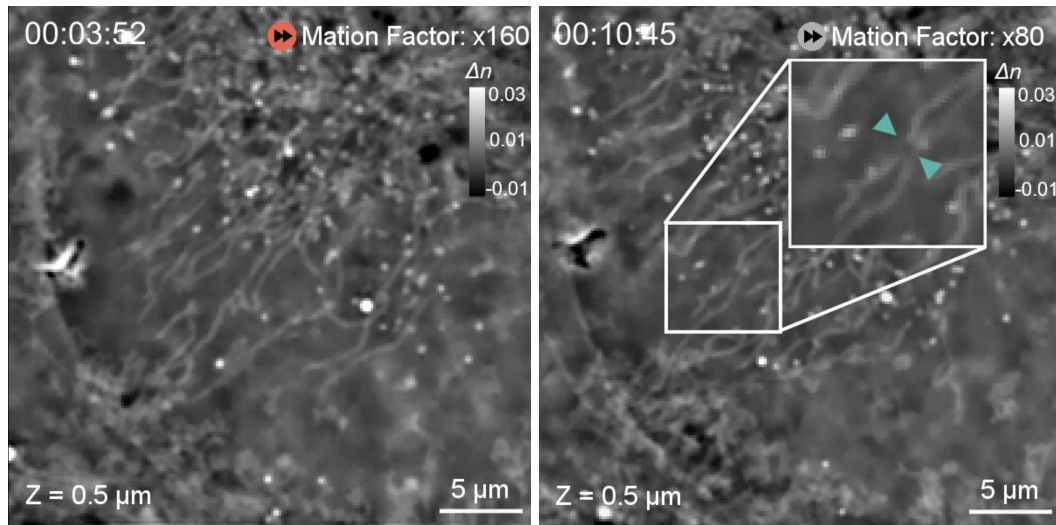
**Supplementary Fig. S12. Visualization of the cell migration process over two hours. a,** Reconstructed RI distribution of  $-0.5 \mu\text{m}$  z plane at four different time points, visualizes the process of cell migration. Both nucleolus in the field of view are located and tracked simultaneously. **b,** RI tomograms at  $0.81 \mu\text{m}$  z plane within the green square box in (a), containing two nuclei with continuous motion over time. The enlarged orange rectangular area is shown in second row over various time points, while the corresponding schematic diagram highlights the interaction between nuclear membrane (orange dotted line), mitochondria (orange), and lipid droplets (yellow) at the time point of 01:38:49. **c,** Another subregion of the RI tomogram in (a) at  $1.14 \mu\text{m}$  z plane shows the changes in cell boundaries and lipid droplets during cell migration (see also Supplementary Video 7). Scale bars: (a)  $20 \mu\text{m}$ , and (b,c)  $10 \mu\text{m}$ .



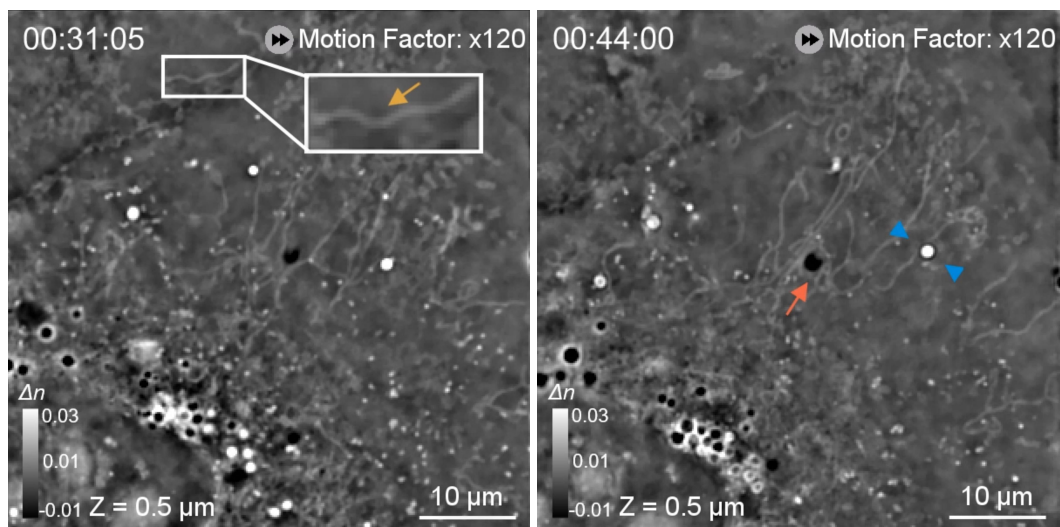
**Supplementary Video S1. Visualization of structures and dynamics associated with pinocytosis and exocytosis.**



**Supplementary Video S2. Visualization of structures and dynamics associated with actin filaments contraction.**

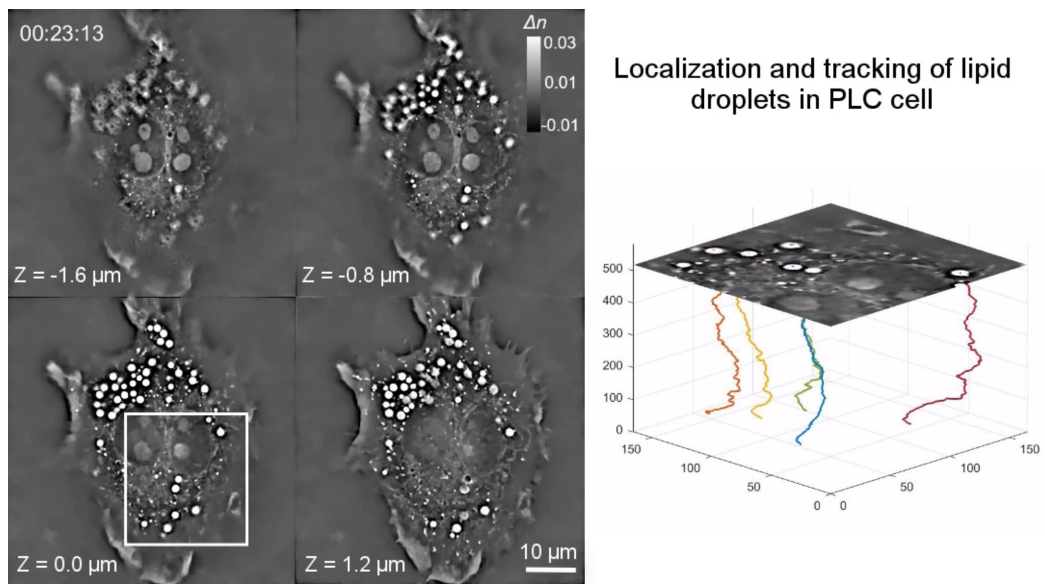


**Supplementary Video S3. Visualization of the mitochondrial fission.**

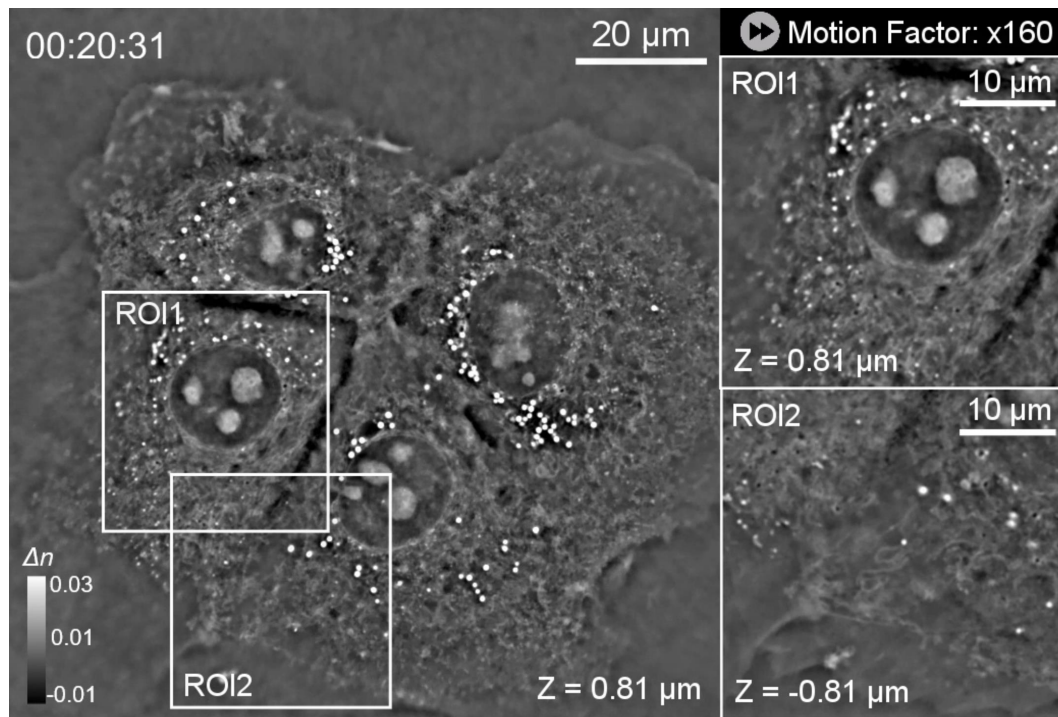


**Supplementary Video S4. Visualization of the mitochondria interacting with other organelles.**

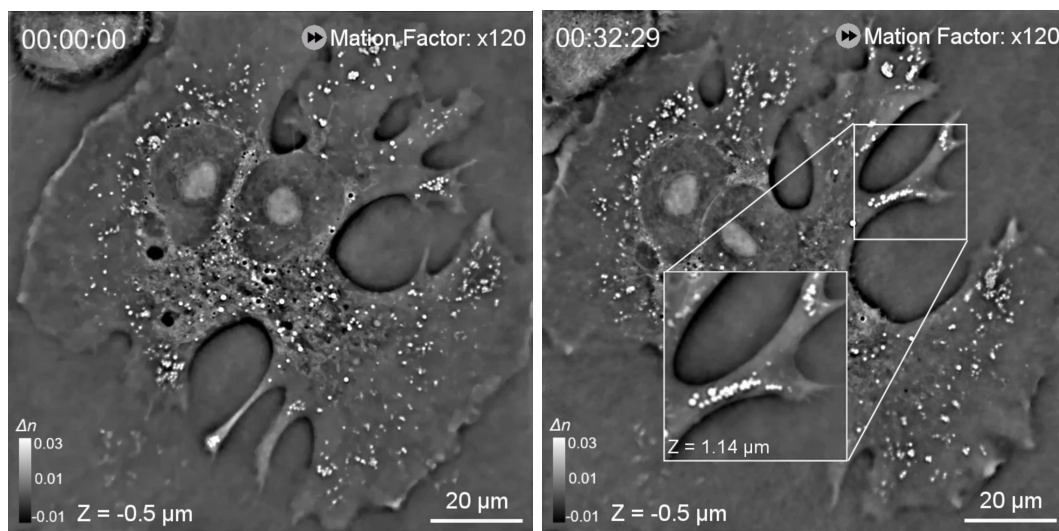




**Supplementary Video S5. Visualization of lipid droplets localization and tracking during cell movement.**



**Supplementary Video S6. Visualization of intercellular interactions during cell growth.**



**Supplementary Video S7. Visualization of cell migration.**